



Article

Simulating and Forecasting the COVID-19 Spread in a U.S. Metropolitan Region with a Spatial SEIR Model

Faizeh Hatami ¹, Shi Chen ^{2,3}, Rajib Paul ^{2,3} and Jean-Claude Thill ^{1,3,*}

¹ Department of Geography and Earth Sciences, University of North Carolina at Charlotte, Charlotte, NC 28223, USA

² Department of Public Health Sciences, University of North Carolina at Charlotte, Charlotte, NC 28223, USA

³ School of Data Science, University of North Carolina at Charlotte, Charlotte, NC 28223, USA

* Correspondence: jfthill@uncc.edu

Abstract: The global COVID-19 pandemic has taken a heavy toll on health, social, and economic costs since the end of 2019. Predicting the spread of a pandemic is essential to developing effective intervention policies. Since the beginning of this pandemic, many models have been developed to predict its pathways. However, the majority of these models assume homogeneous dynamics over the geographic space, while the pandemic exhibits substantial spatial heterogeneity. In addition, spatial interaction among territorial entities and variations in their magnitude impact the pandemic dynamics. In this study, we used a spatial extension of the SEIR-type epidemiological model to simulate and predict the 4-week number of COVID-19 cases in the Charlotte–Concord–Gastonia Metropolitan Statistical Area (MSA), USA. We incorporated a variety of covariates, including mobility, pharmaceutical, and non-pharmaceutical interventions, demographics, and weather data to improve the model’s predictive performance. We predicted the number of COVID-19 cases for up to four weeks in the 10 counties of the studied MSA simultaneously over the time period 29 March 2020 to 13 March 2021, and compared the results with the reported number of cases using the root-mean-squared error (RMSE) metric. Our results highlight the importance of spatial heterogeneity and spatial interactions among locations in COVID-19 pandemic modeling.

Keywords: SEIR model; spatial SEIR model; approximate Bayesian computation; mobility; spatial dependence; temporal variability; epidemic model

Citation: Hatami, F.; Chen, S.; Paul, R.; Thill, J.-C. Simulating and Forecasting the COVID-19 Spread in a U.S. Metropolitan Region with a Spatial SEIR Model. *Int. J. Environ. Res. Public Health* **2022**, *19*, 15771. <https://doi.org/10.3390/ijerph192315771>

Academic Editors: Yibiao Zhou, Jagdish Khubchandani and Paul B. Tchounwou

Received: 7 October 2022

Accepted: 24 November 2022

Published: 27 November 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 contribution of spatio-temporal analysis to the understanding of COVID-19 infections and deaths is by now well documented [1–4]. Not all regions have been impacted with the same rates of infections and deaths. Striking temporal variations were also recorded at both national and regional levels. In early 2020, New York was the epicenter for COVID-19 infections in the US, and then that shifted gradually to Southern states around the July 4th weekend; as Fall approached, the intensity of the disease moved to the Dakotas, and lastly, in January 2022 the Northeast experienced the highest number of new cases of the entire pandemic [5]. The global spread of the disease is linked to international and transregional travel and to different variants of the virus. Local factors such as policies on non-pharmaceutical interventions (e.g., social distancing, mask mandates, school and business closures, etc.), population density and movement patterns, propensity to travel internationally, vaccine availability and uptake rates, and quality of the healthcare infrastructure feature among the core factors of local variations in the spread of the disease and in its prevalence. It is because these local factors condition community transmission, risk of exposure, and community vulnerability to the virus that they are so important at

the local level, and are then amplified across the globe through complex patterns of human spatial interactions.

Given the importance of localized scales, the objective of this study was to simulate and predict the spread of COVID-19 in a multi-county metropolitan region with an epidemic model that explicitly accounts for the spatial and temporal heterogeneities of disease spread. Predictions are made over a four-week forecasting horizon and compared with a series of other forecasting models. Four-week ahead predictions are a common practice in the literature on COVID-19 predictive analytics. For instance, the U.S. Centers for Disease Control and Prevention (CDC) report four-week ahead forecasts from a number of leading infectious disease teams around the world [6].

The spatio-temporal analysis of infectious diseases can be broadly classified into three main groups. In one approach, disease clusters and hotspots are assessed using ecological modeling [2,3]; in the second approach, we study the mechanisms in which the hosts and their movement in space and time contribute to disease transmission; the third technique resorts to machine learning and deep learning methods to tease out the role of space and time in predicting the spread of the disease [7,8]. Ecological models are used often on aggregated data to understand the spatial and spatio-temporal variations in disease spread in terms of prevalence and incidence rates with respect to explanatory variables [9]. These models should not be used for causal inference. Mechanistic models have been extensively applied to characterize various epidemiological processes (e.g., transmission, recovery, vaccination, disease-induced death, etc.), parameterize these critical processes, and provide projections of disease dynamics [10,11]. In particular, mechanistic compartmental models are usually expressed as an array of coupled ordinary differential equations (ODEs), Partial differential equations (PDEs), or stochastic differential equations (SDEs) of various processes [12]. The most commonly applied compartmental model is the susceptible–exposed–infectious–recovered (SEIR) model, with the four compartments representing the host population’s transitioning epidemiological states. The susceptible population is the proportion of the entire population that is capable of contracting the pathogen. The exposed compartment is the population that has been exposed to the pathogen but is not yet infectious. Infectious populations are those who have been exposed to the pathogen and have the capability to further transmit the pathogen. Finally, recovered populations are those who are not infectious any more as a result of recovery with immunity or mortality [12]. Over the course of almost a century, SEIR-type compartmental models have been the paradigm of infectious disease modeling. Conducting a systematic review of 72 publications on COVID-19 predictive models, Shankar et al. [13] found that over 70% of models were in the SEIR group. In another review paper, Rahimi et al. [14] reported that SEIR models are in the top three most common COVID-19 outbreak modeling methods. SEIR-type models have been commonly applied in modeling COVID-19 epidemics and in capturing its complexities, incorporating external factors such as non-pharmaceutical interventions (NPIs) and vaccination [15]. The SEIR-type compartment model is relatively easy to understand, formulate, analyze, and extend to accommodate a wide range of disease-related processes. Important metrics such as the basic reproduction number (R_0) and the effective reproduction number (R_e) can be derived analytically or numerically from the compartment model to evaluate the severity of the pandemic and the effectiveness of interventions, and attempt to forecast the future behavior of the disease.

Modeling frameworks incorporating machine learning or deep learning method leverage the flexibility of these analytical techniques to predict spatio-temporal dynamics. Istaiteh et al. [16] used four machine learning models, including convolutional neural network (CNN), artificial neural network (ANN), long short-term memory (LSTM) and autoregressive integrated moving average (ARIMA) to forecast COVID-19 cases. This study provided forecast for cases over the next 7 days for all countries worldwide. Applying three machine learning methods of hidden Markov chain model (HMM), hierarchical Bayes model, and long short-term-memory model (LSTM), Tian et al. [17] predicted

COVID-19 cases in six countries and used the root-mean-square error (RMSE) to evaluate their prediction performance. Nikparvar et al. [18] used LSTM, a type of recurrent neural network, to predict the number of COVID-19 cases and deaths for 4 weeks and for the entire United States at the fine resolution of counties. Zeroual et al. [19] used five deep learning methods to predict COVID-19 new and recovered cases in six countries. They used recurrent neural networks (RNNs) including LSTM, bidirectional LSTM (BiLSTM), gated recurrent units (GRUs), and variational autoencoder (VAE).

Since SEIR-type models are based on ODEs, they usually do not handle spatial processes (e.g., host movement across space) and spatial heterogeneity (e.g., varying disease-related parameters across space such as different transmission coefficients, rates of recovery and death, etc.) well. Nevertheless, today's human movement patterns are strikingly distinct and more prominent than a century ago when the SEIR model was originally developed. Failure to incorporate modern movement patterns based on spatial heterogeneities may result in unrealistic predictions [20]. Therefore, new methodologies, such as spatial SEIR models, have been developed to address and accommodate the critical spatial component in infectious disease dynamics explicitly. Several new SEIR-based models have recently been advanced that alleviate the limitations of earlier generations of SEIR models to be more in tune with the patent realities of highly mobile and heterogenous contemporary societies. This is the approach we followed in this study.

Spatial SEIR models generally incorporate two or more spatially connected and interacting populations, each having their own parameters related to disease processes (e.g., within-population transmission rate, recovery rate, death rate in the specific population, etc.). Within each population, the disease dynamics can be characterized by the traditional SEIR model as if the population is isolated (e.g., no interactions among populations). More importantly, these populations have interactions (inflow and outflow via commuting) that lead to new transmissions. Intensity or rate of interactions will influence potential transmission, thus changing the overall disease transmission dynamics within and among populations.

In this study, we used a spatial SEIR model to simulate the number of COVID-19 cases and forecast their four-week ahead values in ten counties of the Charlotte–Concord–Gastonia Metropolitan Statistical Area (MSA). In this model, the number of cases was forecasted for all locations in one model, simultaneously. The spatial interactions between different locations were incorporated into the model to enhance the model's accuracy of simulations and predictions. In addition, due to the impacts of external factors on COVID-19 transition dynamics, a set of covariates were included in this model, including the sociodemographic factors, mobility, county-level political leaning, COVID-19 vaccination coverage, and intervention policies. The model's predictive performance is evaluated by computing the root-mean-square error (RMSE) between model predictions and reported values.

2. Materials and Methods

2.1. Study Area

We studied the spatio-temporal dynamics of COVID-19 in the Charlotte–Concord–Gastonia Metropolitan Statistical Area (MSA). This region is one of the largest MSAs in the Southeastern U.S. It comprises a total of 10 counties (7 in North Carolina and 3 in South Carolina) with an estimated total population over 2.6 million in 2021 [21] (Figure 1). Its population has grown by 16% from 2010 to 2021; during this period, population growth reached 22% in Mecklenburg County, North Carolina's second most populous county [22]. As the most important commuting destination in the MSA, Mecklenburg County has experienced 32% of growth in employment in 10 years (from 697,231 in 2010 to 923,259 in 2020) [23]. It constitutes the employment center of the MSA as about 60% of commuting trips end in this county [24]. Other counties trail Mecklenburg County both in population

and employment, and operationally have a residential function within the MSA interlaced with a few outlying business centers.

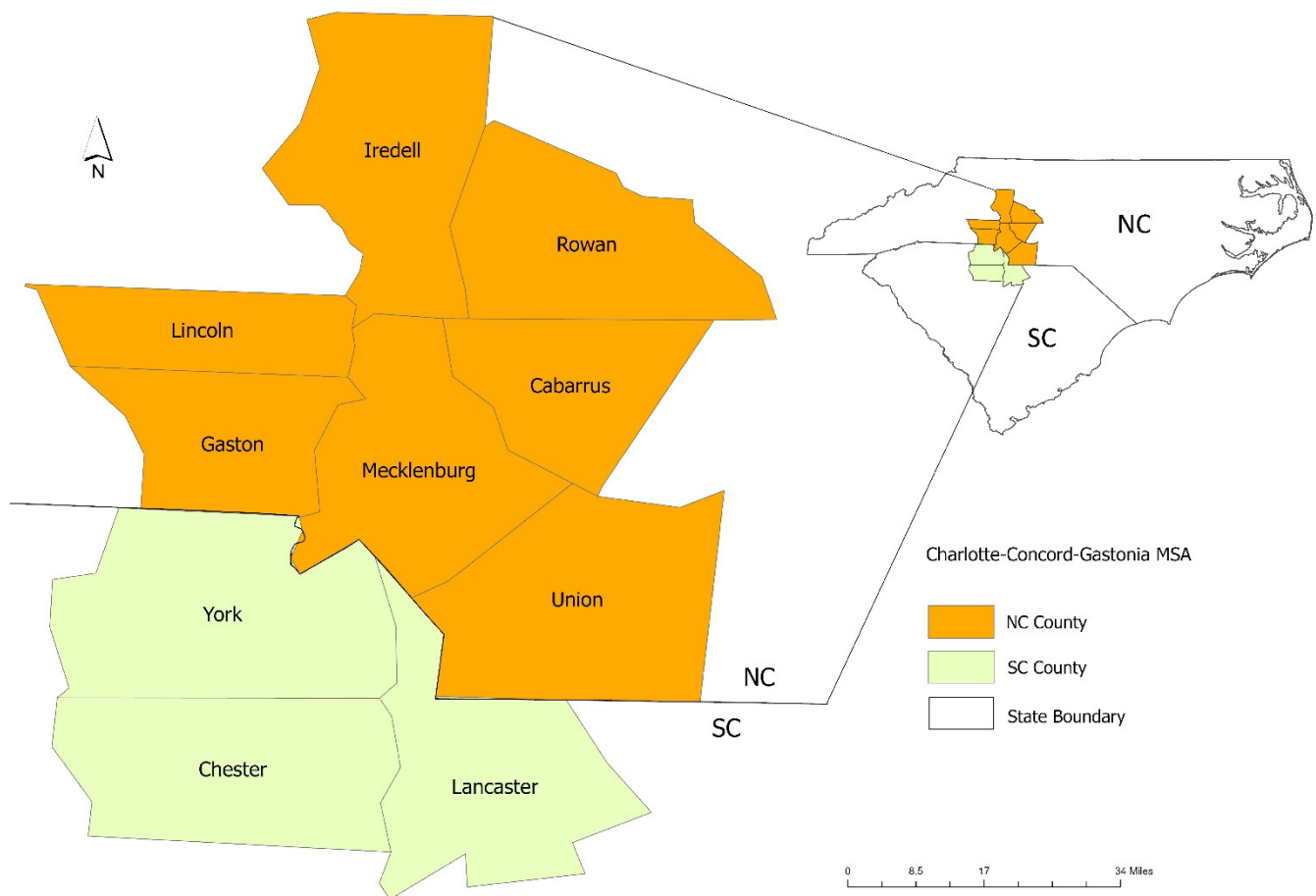


Figure 1. Map of the Charlotte–Concord–Gastonia MSA.

In addition to fast population and economic growth, the Charlotte area has been experiencing transformative urban development and growth in two major forms relevant to the spread of COVID-19 pandemic: (1) high-intensity suburban developments in the outer parts of the city; and (2) high density and mixed developments in the inner parts of the city. Compact urban developments in the inner parts include the vertical developments in the central areas, infilling in the older neighborhoods around the central city and transit-oriented developments along the transit corridors [25]. Land use transportation plans have been developed to integrate the land developments with a rapid transit system along transit corridors. In addition to the extensive highway system, the LYNX light rail service has enhanced connectivity between different parts of the county and beyond [26].

The county's large population size, rapid growth with high socio-economic disparity, fluid human movement, and its dominance of the entire MSA make this region vulnerable to large epidemics such as the current COVID-19.

2.2. Data Retrieval and Preprocessing

Data for the cumulative counts of COVID-19 cases and deaths in a daily temporal resolution were sourced from the Center for Systems Science and Engineering (CSSE) at Johns Hopkins University [27]. Data were collected from 29 March 2020 to 13 March 2021 for each of the 10 counties in our study area, and then cleaned and preprocessed for fitting the spatial SEIR prediction model (Section 2.3). Data for the following four weeks were used for prediction purposes based on the fitted model. We restricted our analysis to this

study period mainly for three reasons: (1) We wanted to study one type of variant of COVID-19, Delta variant (B.1.617.2) started interfering around 20 May 2021, and later became the dominant variant. Thus, we restricted our study prior to arrival of Delta variant; (2) Data monitoring in the regions under study gradually declined as severity of the cases declined and vaccines became more widely available; (3) Finally, beyond 31 March 2021, all nonpharmaceutical restrictions in the region were lifted. During data preprocessing, it was found that some cumulative cases and deaths had smaller values than on the previous day. This issue is resolved by replacing the incorrect value by the value of the previous day, thus ensuring monotonically increasing cumulative cases and deaths. Daily incident cases and new deaths were obtained from the cleaned daily cumulative data through differencing values between two consecutive days. Then, 7-day moving averages of new daily incident cases and new deaths were calculated using a centered moving average method. Finally, daily moving averages of incident cases and new deaths were aggregated to weekly counts. Weekly aggregates were consistent with the Morbidity and Mortality Weekly Report (MMWR) weeks [28].

In addition to the COVID-19 case and death data, covariates critical to COVID-19 dynamics were collected and incorporated in the prediction model: sociodemographic factors such as population density, age and race, mobility, county-level political leaning, COVID-19 vaccination coverage, and intervention policies. These covariates were structured in two groups of time-variant and time-invariant variables, respectively. Time-invariant data were used to capture spatial heterogeneity in this study (i.e., differences between spatial entities, counties), while time-variant variables were used to capture spatio-temporal heterogeneity. In addition, time-variant variables were modeled with a 2-week lag to account for the COVID-19 incubation period [29].

Sociodemographic factors of race, population density, and age distribution were sourced from the SimplyAnalytics database [30], drawing primary data from the US Census Bureau American Community Survey. Mobility patterns are one of the most important covariates of the COVID-19 pandemic [18,31,32]. In order to capture the impact of the interactions between populations in different counties, mobility data were collected from different sources and incorporated in the prediction model. The daily time spent at different places such as home, work, public transit, grocery stores, and parks was obtained from Google COVID-19 Community Mobility Reports [33]. Missing data were imputed by values of the previous day. Daily values were aggregated to weekly values based on the same epidemiological calendar of MMWR. We also used Apple Mobility [34] data, which include the percentage change in mobility on a daily basis, as a reference to cross-validate other mobility data sources. Apple Mobility daily values were aggregated to weekly values for counties following the same approach mentioned earlier. In addition, we collected the total number of individuals' visits to points of interest (POIs) from the SafeGraph's Monthly Patterns dataset [35]. SafeGraph collects mobility data from a variety of resources such as mobile phone GPS and provides these data through the SafeGraph COVID-19 Data Consortium [36]. In this dataset, the total tally of daily visits to POIs is available by Census Block Group. In this study, Census Block Group level data were aggregated to counties, and daily visits were aggregated to weekly visits for each county based on the epidemiological calendar.

Another important societal factor that is able to influence perception of the risk of COVID-19 and behavioral responses (e.g., whether or not one is willing to comply with various interventions) is the political leaning of the places [37]. For this purpose, we captured the differences between counties in the voting results of the 2020 presidential elections. Data were collected from the North Carolina State Board of Elections [38] for North Carolina counties, and from the South Carolina Election Commissions [39] for South Carolina counties. Then, the ratio of votes for the Democratic candidate over the Republican candidate was calculated and incorporated into the prediction model.

Given the importance of COVID-19 vaccination, especially its impact on the third and largest wave in COVID-19 cases in 2021, county-level vaccination coverage with at least

one dose was obtained for North Carolina counties from the North Carolina Department of Health and Human Services [40]. Historical vaccination data are available for North Carolina counties; however, for South Carolina, only the cumulative data are available. As a result, the weekly counts of vaccinated people with at least one dose for SC counties were estimated on the population-weighted vaccination counts in NC counties. Then, the weekly counts were aggregated to cumulative values to account for the continuous impact of vaccinations over time.

In addition, several other policy-related variables were used; for example, the mandates on face covering starting in June 2020 [41], school shutdowns starting in August 2020. Wintery atmospheric temperature between October 2020 and February 2021 is used as an environmental factor. Descriptive statistics of important variables are summarized in Table 1.

Table 1. Descriptive statistics of variables used in the final exposure model.

Variable	Mean	Std. Dev.	Min	Pctl. 25	Pctl. 75	Max
Google mobility (change in time spent at workplaces (%))	-28.835	8.118	-53.714	-33.893	-22.714	-10.429
Population vaccinated with at least one dose	1028.341	2375.554	0	0	306.25	13162
Apple mobility (average requests for changing directions)	130.352	24.699	53.999	119.248	148.376	180.517
Voting ratio (democrats/republicans)	0.756	0.468	0.361	0.508	0.8	2.094
Population density (per mi2)	515.209	551.934	55.335	273.888	579.502	2097.705
Wintery temperature	0.63	0.48	0	0	1	1
School shutdowns	0.39	0.49	0	0	1	1
Face mask intervention policy	0.28	0.45	0	0	1	1

The preprocessed COVID-19 cases were used to construct the data model and the mentioned covariates are used to build the exposure model component of the overall spatial SEIR model in this study (detailed in Section 2.3). Another important component of this spatial model is the distance model (also known as the spatial process model), which specifies exposure probabilities as a linear combination of so-called ‘spatial weight matrices’. These matrices capture the spatial structure of the units of observation; specifically, they measure the geographic proximity (or conversely, their geographic separation of distance) on a pairwise basis. In this study, they were based on the length of the shared boundary between counties and on the inverse of the pairwise distance between the centroids of all 10 counties, using the Cartographic Boundary Files-Shapefile [42]. The preliminary analysis led to the conclusion that such a combination of spatial structures provided a more comprehensive characterization of spatial interaction effects and a more effective handling of disease diffusion across the geographic space than either of these two matrices alone. The two spatial weight matrices were unweighted in the distance model.

2.3. Modeling Framework and Model Structure

Brown et al. [43] developed a stochastic spatial SEIR model fitted to data using the approximate Bayesian computational algorithm [44]. It models the dynamics of the epidemic over time at multiple spatial locations. Equation (1) shows the mathematical formulation of the model with four compartments (S, E, I, R), and transitions into new compartments (S*, E*, I* and R*). New transitions indicated by an asterisk occur over discrete times (*i*) and discrete locations (*j*). $\pi_{ij}^{(RS)}$, $\pi_{ij}^{(SE)}$, $\pi_{ij}^{(EI)}$ and $\pi_{ij}^{(IR)}$ show the probabilities of transitions between R and S, S and E, E and I, and I and R, respectively.

$$\begin{aligned}
S_{i+1} &= S_i - E_i^* + S_i^* & S_{ij}^* &\sim \text{binom}(R_{ij}, \pi_{ij}^{(RS)}) \\
E_{i+1} &= E_i - I_i^* + E_i^* & E_{ij}^* &\sim \text{binom}(S_{ij}, \pi_{ij}^{(SE)}) \\
I_{i+1} &= I_i - R_i^* + I_i^* & I_{ij}^* &\sim \text{binom}(E_{ij}, \pi_{ij}^{(EI)}) \\
R_{i+1} &= R_i - S_i^* - R_i^* & R_{ij}^* &\sim \text{binom}(I_{ij}, \pi_{ij}^{(IR)})
\end{aligned} \tag{1}$$

where the exposure probabilities $\{\pi_{ij}^{(SE)}\}$ are modeled using exposure-related covariates (such as mobility, vaccinations, population demographics, non-pharmaceutical interventions (NPIs) and weather) and spatial weight structures. The spatial structure of the population is incorporated into the exposure probability as shown in Equation 2. In this equation, D_z denotes n-by-n spatial weight matrices, and P denotes the associated autocorrelation parameters.

$$\pi_{ij}^{(SE)} = 1 - \exp\left(-\eta_i - \sum_{z=1}^Z P_z(D_z \eta_i)\right)^{h_i} \tag{2}$$

$\pi_{ij}^{(SE)}$ in Equation (2) indicates the probability of transitioning a susceptible individual in location s_j at time t_i into the exposed population.

The probabilities of transitioning from E to I, and from I to R, are explained in Equation (3), in which γ denotes the transition ratio parameter and h_i denotes a temporal offset allowing for irregularly spaced time points. More detailed descriptions of the model can be found in Brown et al. [43].

$$\begin{aligned}
\pi_{i}^{(EI)} &= 1 - \exp(-h_i \gamma_{(EI)}) \\
\pi_{i}^{(IR)} &= 1 - \exp(-h_i \gamma_{(IR)})
\end{aligned} \tag{3}$$

The ABSEIR R package [43] was used to simulate the SEIR model. As a stochastic spatial SEIR model, the overall model is composed of several modules: data model, exposure model, reinfection model, distance model, transition priors, initial value container, and sampling control. After each individual module has been developed, all modules are fed into the spatial SEIR model to simulate the number of reported case series. The data model component describes the input data and their relationship with the SEIR model. A matrix of COVID-19 cases with rows denoting the temporal dimension and columns denoting different locations is fed into the data model. In addition, other parameters such as the matrix type (“identity”, “overdispersion”, or “fractional”), and compartment (I^* or R^*), were included in the data model. In the data model, it needs to be specified whether the input data (number of cases) are in cumulative format or indicate new cases in each time point. The exposure model component describes the actual epidemic process including various covariates. The reinfection model component indicates whether transitions from recovered (R) to susceptible (S) compartments are allowed in the model or not. The transition priors component specifies the prior transition probabilities from E to I, and from I to R compartments. These two transition priors capture the latency period and infectious period of the dynamics. The distance model describes the spatial network structure of the population as it conditions the interactions among counties in the MSA. The distance model is composed of a list of square, symmetric spatial weight matrices. Each square matrix indicates the pairwise relationships between studied locations. The pairwise relationship can be defined in different ways. For example, it can show the straight-line distance between location centroids, or whether two locations are neighbors or not. The initial value container is constructed based on the total population, number of cases and deaths, and COVID-19 tests at the beginning of the study period, in each county. Lastly, sampling control specifies the approximate Bayesian computation (ABC) algorithm and other sampling properties. Details of the model can be found in the ABSEIR GitHub repository (<https://github.com/grantbrown/ABSEIR>, accessed on 15 November 2020) by Brown et al. [43].

Once the complete model is constructed and fitted, COVID-19 cases are predicted for up to four weeks ahead for dates between 20 March 2021 to 10 April 2021. Then, model performance (goodness of fit) is evaluated by computing the root-mean-square error (RMSE) between model predictions and reported values. Note that each of the 10 counties in this study will have its own, county-specific model parameters and RMSEs. We developed different versions of the model by including different specifications of covariates (i.e., inclusion of certain combinations of covariates), and optimal models are identified by comparing RMSEs of competing models at each county. The stability of our models' performance is tested by running the same models on another 4-week period from 17 April 2021 to 8 May 2021. The results of this 4-week period were also compared with the ensemble model as a benchmark.

In order to assess the enhancement made by the inclusion of the spatial interactions' characteristics in the SEIR model, we also developed comparable models without the spatial interactions. These nonspatial models simultaneously estimate COVID-19 cases for all counties without taking into account their geographic connections. In these models, counties were assumed to be isolated from each other (spatial weight matrices were null).

In addition, we further evaluated the performance of our fitted spatial SEIR models by comparison with models developed in the extant literature. Specifically, we compared our model outputs with an ensemble prediction model implemented by Reich Lab's COVID-19 Forecast Hub in collaboration with CDC [45]. This ensemble model takes the predicted values of COVID-19 cases and deaths from over 50 international research groups on a weekly basis, and calculates the arithmetic mean and median to construct the ensemble. The COVID-19 Forecast Hub's ensemble model is a good benchmark to evaluate the performance of our model. The predictive performance of our model on the alternate period from 17 April 2021 to 8 May 2021 was also compared with the ensemble model as a benchmark.

3. Results

Multiple spatial SEIR models were constructed based on different specifications for the exposure model component, i.e., with different combinations of covariates/factors. Only the prediction results of three models with best goodness-of-fit measures are reported here. The covariates in each of these models are listed in Table 2.

Table 2. Three retained spatial SEIR models: specification and approximate Bayes factors.

	Model Specification	Approximate Bayes Factor		
		Model 1	Model 2	Model 3
Model 1	Face mask intervention policy	1.0	0.2	3.1
	School shutdowns			
	Wintery temperature			
	Population density			
	Vaccinated population with at least one dose			
Model 2	Face mask intervention policy	5.6	1.0	17.3
	School shutdowns			
	Wintery temperature			
	Presidential election voting ratio			
	Vaccinated population with at least one dose			
	Percentage change in mobility (Apple Mobility dataset)			
Model 3	Face mask intervention policy	0.3	0.1	1.0
	School shutdowns			

Wintery temperature
Presidential election voting ratio
Vaccinated population with at least
one dose
Change in time spent at work (Google
mobility reports dataset)

The three models are reported as each of them performed better than the others in some of the counties. In the other words, no single model performs consistently well across all 10 counties in this study. The spatial variability in model performance indicates that particular spatial and/or spatio-temporal covariates impact COVID-19 dynamics differently in different locations.

Bayes factor, the ratio of the acceptance between two models, is used to compare the overall performance of the three optimal models [46]. As reported in Table 2, all values of the Bayes factor are greater than 3 in model 2, indicating moderate to strong evidence for model 2's overall better performance than the other two models. On the other hand, none of the factors exceeds 3 in model 3, and only one exceeds 3 in model 1. Taken together, these results suggest that model 2 outperforms the other models. Figure 2 shows fitting and prediction results of model 2 as the overall best model.

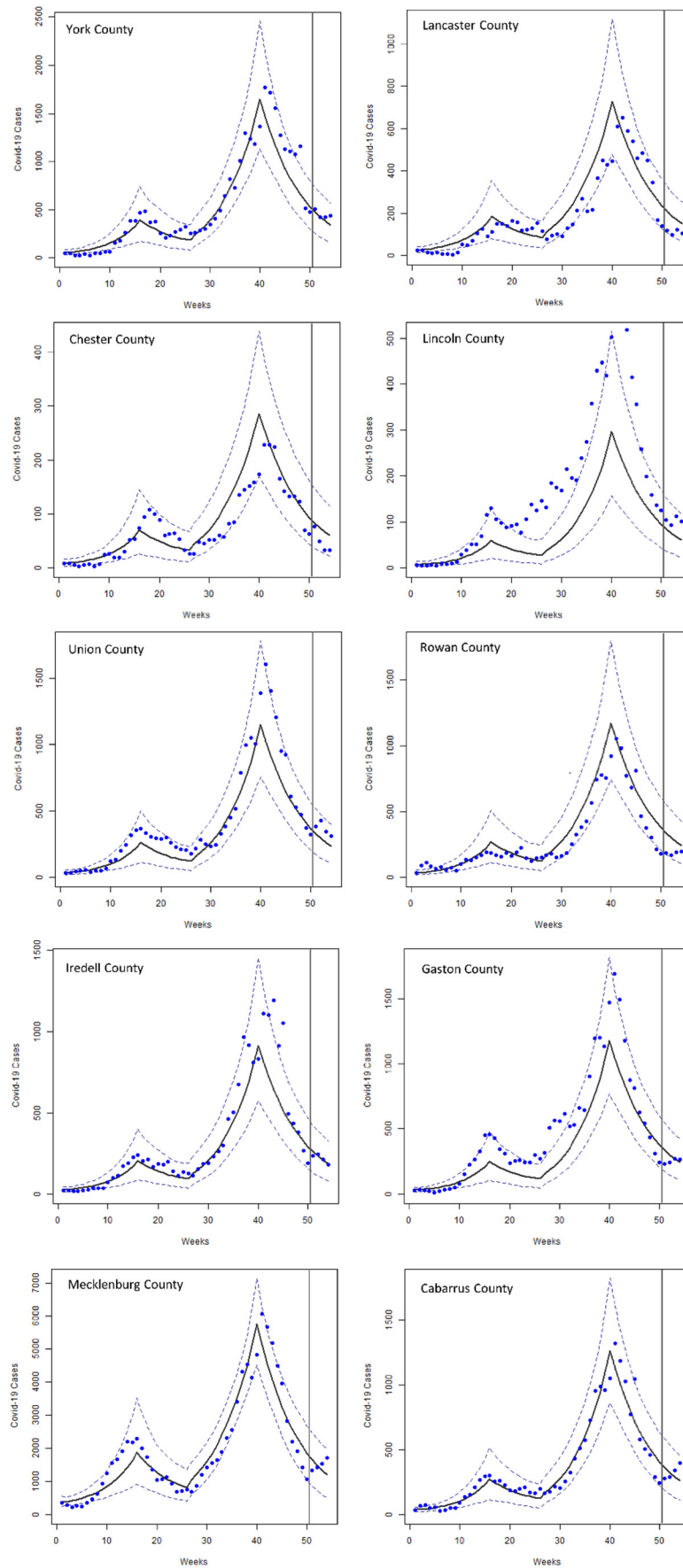


Figure 2. Reported historical (blue dots) COVID-19 cases and cases simulated with model 2 (black line) in 10 counties of the study area, including the last 4 weeks not used for model fitting. The 5% confidence interval is shown with dashed lines. The last 4 weeks to the right of the gray vertical lines show the forecasted number of cases.

Table 3 shows the 4-week ahead prediction of COVID-19 cases with each model between 20 March 2021 to 10 April 2021, as well as the reported cases for benchmarking purposes. With this information, the RMSE values can be calculated for our spatial SEIR model and similarly for the ensemble model from CDC [45]. These are reported in Table 4. Using the ensemble model as a benchmark, we find that our spatial SEIR model performs better than the benchmark model in 8 of the 10 counties (Cabarrus, Gaston, Iredell, Mecklenburg, Union, Rowan, Lancaster and York), weaker in just one county (Lincoln), and on par in Chester County. To compare the overall predictive performance of each model in all counties with the overall performance of the benchmark ensemble model, one overall weighted RMSE value is calculated for each model across all counties and weighted by the county population size (Table 5). We also calculated an overall unweighted RMSE. All of the spatial SEIR models developed in this study perform drastically better in predicting 4-week ahead cases than the benchmark ensemble across all counties.

Table 3. Four-week predicted cases (PC) and reported cases (RC) across 10 counties in the MSA.

		County									
		Cabarrus	Gaston	Iredell	Lincoln	Mecklenburg	Rowan	Union	Chester	Lancaster	York
PC Model 1	W1	284	251	226	64	1677	265	299	72	207	454
	W2	256	227	204	58	1515	240	270	65	188	410
	W3	232	205	184	53	1372	217	244	59	170	370
	W4	210	186	167	47	1241	196	221	53	153	336
PC Model 2	W1	364	337	261	85	1665	338	328	84	209	472
	W2	326	301	234	76	1492	303	294	75	188	423
	W3	292	270	210	68	1340	271	263	67	168	379
	W4	262	243	188	61	1203	243	236	60	151	340
PC Model 3	W1	181	170	132	36	1661	160	174	41	113	269
	W2	162	151	118	32	1484	143	156	37	101	240
	W3	144	135	105	29	1328	127	139	33	90	214
	W4	130	121	94	26	1189	114	125	29	81	192
RC	W1	277	233	236	104	1328	184	383	77	120	503
	W2	289	245	244	93	1436	166	424	50	100	419
	W3	343	270	213	112	1532	190	346	33	123	420
	W4	396	268	183	101	1724	196	313	32	106	440

Table 4. Four-week ahead prediction RMSE values of the three variants of the spatial SEIR model and of the ensemble model for each county.

County	Cabarrus	Gaston	Iredell	Lincoln	Mecklenburg	Rowan	Union	Chester	Lancaster	York
Spatial SEIR—Model 1	110	54	26	48	311	56	111	18	70	63
Spatial SEIR—Model 2	86	60	14	32	326	113	90	26	70	56
Spatial SEIR—Model 3	184	115	108	72	332	54	220	19	21	218
Ensemble	174	92	102	23	882	86	148	18	22	139

Table 5. RMSE values of prediction models over 4 weeks and all counties.

	Model 1	Model 2	Model 3	Ensemble
RMSE- average weighted by county population	172	177	224	440
RMSE- unweighted average	87	87	134	169

The three spatial models are also compared with their nonspatial counterparts to establish the specific contribution of the spatial dimension on the modeling of the spread of COVID-19. Integrating the spatial and spatio-temporal dynamics across counties has substantially improved the predictive performances of all specifications of the SEIR model. Specifically, we find that the average 4-week RMSE across all 10 counties has improved by 47%, 31% and 80% in models 1, 2 and 3, respectively, compared with the nonspatial counterparts. There is latent spatial heterogeneity in the spread of COVID-19, which transpires through the predictive performance of the instances of spatial SEIR model as no specification performs better than the others in all 10 counties. Model 1 performs better than other models in three counties; model 2 better than other models in five counties, and model 3 better than other models in two counties. Hence, the overall RMSE values (Table 5) largely hide these differences. However, we note that the weighted RMSE identifies model 1 as the best performer due to the overwhelming population size of Mecklenburg County over other counties, while the unweighted RMSE value puts models 1 and 2 on an equal footing.

Finally, we test the stability of the predictive performance of models over time. For this purpose, we run the same model specifications for another 4 weeks (from 17 April 2021 to 8 May 2021). We have found that our fitted models perform consistently well across prediction periods compared with the benchmark ensemble model.

4. Discussion

We have developed a spatial SEIR model to fit the COVID-19 epidemic dynamics and implemented it to predict the number of 4-week ahead cases. SEIR models have been one of the most commonly used methods for studying epidemic dynamics. However, original SEIR models have no explicit spatial component, and cannot adequately capture potential spatial processes and variability across different locations. Spatial SEIR, on the other hand, takes into account the heterogeneity of epidemics across different locations. It also captures the spatial interactions (e.g., mixing via movement and mobility) among these locations.

The spread of COVID-19 is simulated and predicted for future weeks in the 10 counties of the Charlotte–Concord–Gastonia MSA. The relative positions of Mecklenburg County and of the surrounding counties in the MSA, and their respective social determinants of health (e.g., employment rates) and landscape structure in this broader geographic context, are generative of meaningful spatial heterogeneities and lead to strong interactions between these counties including commuting trips, non-work trips, and mobility of goods and services.

Metropolitan areas have been particularly impacted during the COVID-19 pandemic because of spatial interactions and heterogeneity within and between locations in the metropolitan areas. According to the results of this study, the spatial SEIR model proves effective at accounting for these complicated spatial interactions and variations to forecast the future COVID-19 epidemic dynamics. Our model outperforms similarly specified SEIR models by a large margin because it explicitly accounts for spatial context and spatial heterogeneity and for the existence of heterogeneous interactions between constituting spatial parts of the study region. Extensive research in spatial analysis of urban and regional systems [47] has considerably advanced the state-of-the-art in this area. The present study points that this is also the case with population health and with the spread of infectious diseases such as the COVID-19 pandemic. The epidemic dynamics in different counties are highly influenced by the spatio-temporal heterogeneity of external covariates. Model performance varies across counties, which indicates the important role of external covariates and of their spatio-temporal heterogeneity in the spread of COVID-19 in a large, heterogeneous metropolitan region. In our research, when the structure of spatial interactions operationalized by two spatial weight matrices was controlled for during the model fitting process (in effect, assuming that counties are compartmentalized and that there are no interactions between them), it was determined that the goodness-of-fit and predictive power of the SEIR model were significantly downgraded. Hence, the structure of the spatial environment matters as far the spatio-temporal pattern of the incidence of COVID-19 is concerned.

Several studies have addressed the spatial patterns and heterogeneities of COVID-19 in SEIR-type models. For example, Hou et al. [48] treated spatial heterogeneities in inter-county modeling of COVID-19 in Wisconsin using business movement patterns, age, and race as covariates. Embedding a network-cluster-based approach in human mobility flow-augmented stochastic SEIR models, they estimated a region-specific R_e . They used an Ensemble Kalman Filter approach to fit their proposed models to the data obtained from the Wisconsin Department of Health Service. In another study, Lawson and Kim [49] used a Bayesian space-time SIR model for estimating and predicting county-level COVID-19 infections and deaths in South Carolina. Region-specific model parameters were modeled using intrinsic conditional autoregressive (ICAR) priors. Chiang et al. [50] used the Hawkes process with spatio-temporal covariates based on county-level Google mobility data and demographics for estimating incidences, deaths, and basic reproduction numbers across the 48 contiguous US states via an expectation-maximization algorithm. Kinsey et al.'s [51] Bucky model used a stratified SEIR model based on age groups and geographic locations to estimate case rates, deaths, and healthcare burden. In a more recent study, emphasizing the geographic heterogeneity of pandemic dynamics, Macias et al. [52] proposed a Lagrangian-SEIR-based model to predict the COVID-19 epidemic in the Mexican state of Jalisco. Their approach incorporated a Lagrangian movement model that captures population movements within and among regions. Gopalakrishnan et al. [53] emphasized the impact of spatial granularity on COVID-19 forecasting results using different unit areas of state, county cluster and county; however, they did not take into account the spatial interactions between these areal units in their compartmental forecasting model. Liu and Li [54] proposed a multi-group SEIR model taking into account the spatial heterogeneity through incorporating spatial diffusion and heterogeneity in the model parameters. In addition, emphasizing on the importance of capturing the spatial interactions in SEIR models, Rajuladevi et al. [55], studied the impact of seeding or initializations in a network of different countries. They found that the effect varies over different countries. Our work expanded this literature by exemplifying the application of spatial SEIR on a metropolitan region and assessed the spatial variations in models' predictive power. Our work opens avenues for future research on fitting various models with different covariates in different regions under spatial SEIR set up with appropriate penalization [56] and variable selection [57,58] techniques. This study has several limitations. While spatial SEIR represents the future of epidemic modeling and forecasting, questions remain to be

investigated to realize the full potential of this sub-family of models. First, it remains to be determined what epidemics benefit to be modeled with an explicit representation of spatial heterogeneity and interactions (also known as spatial dependence). To what extent would different models of space (especially of the spatial weight matrix) for different types of viral infections enhance the performance of the models? Spatial analysis and spatial econometrics have made great strides over the past decade in some respects with a number of modeling frameworks, but not with SEIR-type models to this day.

Second, given the known behaviors of the COVID-19 epidemic, we see the need to develop models at a higher spatial resolution than the current county-level to explicitly represent community-level transmission, and even multi-level models with nested spatial representations that match the range of diffusion processes between place-based populations. Third, we envision that machine learning can greatly enhance the performance of spatial SEIR models by optimizing model parameters that fit the local geographic contexts. In addition, many diseases have concurrent variants/strains such as COVID-19 (e.g., transition period between Alpha and Delta, and Delta and Omicron variants in the U.S.), which substantially influence epidemic dynamics. With broader and more accurate genomic and serosurveillance, we will expand spatial SEIR models to explicitly incorporate multiple variants/strains and their composition in different regions.

5. Conclusions

In this study, we developed a spatial SEIR model capable of simulating and predicting the number of 4-week ahead COVID-19 cases in the Charlotte–Concord–Gastonia MSA. The analysis concluded the spatial SEIR is effective at predicting the spatio-temporal dynamics of the pandemic by explicitly accounting for spatial heterogeneity and spatial interaction between counties.

In addition to pathogen biology, COVID-19 pandemic is influenced by various external covariates such as host mobility, demographic factors, non-pharmaceutical interventions such as school closures, and pharmaceutical interventions such as vaccinations. As a result, the epidemic dynamics vary across different locations based on variations in these factors. In this study, we developed different models incorporating a variety of external covariates and relevant datasets, such as face mask intervention policies, school closings, seasonal atmospheric temperature drops, population density, vaccinated population with at least one dose, presidential election voting ratio, percentage change in mobility from Apple Mobility dataset, and change in time spent at work from the Google mobility reports dataset. We found that spatio-temporal heterogeneity of these external covariates is a strong predictor of the epidemic dynamics. We have identified both the time variant and time invariant covariates that are critical to local (county-specific) COVID-19 epidemic dynamics. Among all competing models that we have developed, three models perform better than others, based on RMSE values as the model evaluation metric. The variation in model performances across counties harks back to the influential impact of external covariates and their spatio-temporal heterogeneity in COVID-19 diffusion in a diverse metropolitan region.

Our analysis has shown that the spatial SEIR model is effective at accounting for the complicated spatial interactions and variations within and between locations in the metropolitan areas to forecast the future COVID-19 epidemic dynamics. Our models outperformed equivalent SEIR models that rule out spatial dependence and spatial interaction between counties. We also compared and evaluated our model with the CDC's ensemble model. Our spatial SEIR models performed on a par with, or even better than, the benchmark ensemble model, showing the importance of incorporating high-resolution, location-specific covariates to further enhance model predictability at the metropolitan scale.

With regard to the tremendous socioeconomic costs of the COVID-19 outbreak, it is a high priority to be able to predict the epidemic dynamics beforehand. This information can be used to inform timely and effective policy such as pharmaceutical and non-pharmaceutical interventions. Our spatial SEIR model developed for Charlotte–Concord–

Gastonia MSA can be readily extended to other regions (especially metropolitan areas with high spatio-temporal heterogeneity) to effectively characterize and predict the COVID-19 epidemic dynamics.

Author Contributions: Conceptualization, R.P., S.C. and J.-C.T.; methodology, F.H., R.P., S.C. and J.-C.T.; validation, F.H., R.P., S.C. and J.-C.T.; formal analysis, F.H.; investigation, F.H., R.P., S.C. and J.-C.T.; data curation, F.H.; writing—original draft preparation, F.H., R.P., S.C. and J.-C.T.; writing—review and editing, F.H., R.P., S.C. and J.-C.T.; visualization, F.H.; supervision, R.P., S.C. and J.-C.T. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Data used in this study will be provided upon request.

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

References

- Franch-Pardo, I.; Napoletano, B.M.; Rosete-Verges, F.; Billa, L. Spatial analysis and GIS in the study of COVID-19. A review. *Sci. Total Environ.* **2020**, *739*, 140033.
- Paul, R.; Adeyemi, O.; Ghosh, S.; Pokhrel, K.; Arif, A.A. Dynamics of COVID-19 mortality and social determinants of health: A spatiotemporal analysis of exceedance probabilities. *Ann. Epidemiol.* **2021**, *62*, 51–58.
- Paul, R.; Arif, A.A.; Adeyemi, O.; Ghosh, S.; Han, D. Progression of COVID-19 from urban to rural areas in the United States: A spatiotemporal analysis of prevalence rates. *J. Rural. Health* **2020**, *36*, 591–601.
- Rahman, M.M.; Thill, J.C. Associations between COVID-19 Pandemic, Lockdown Measures and Human Mobility: Longitudinal Evidence from 86 Countries. *Int. J. Environ. Res. Public Health* **2022**, *19*, 7317. <https://doi.org/10.3390/ijerph19127317>.
- The New York Times. Cases by Region. 2022. Available online: <https://www.nytimes.com/interactive/2021/us/COVID-cases.html> (accessed on 9 February 2022).
- Cramer, E.Y.; Huang, Y.; Wang, Y.; Ray, E.L.; Cornell, M.; Bracher, J.; Brennen, A.; Rivadeneira, A.J.C.; Gerding, A.; House, K.; et al. The United States COVID-19 Forecast Hub dataset. *Sci. Data* **2021**, *9*, 462. <https://doi.org/10.1101/2021.11.04.21265886>. Available online: <https://www.medrxiv.org/content/10.1101/2021.11.04.21265886v1> (accessed on 17 August 2021).
- Lalmuanawma, S.; Hussain, J.; Chhakhuak, L. Applications of machine learning and artificial intelligence for COVID-19 (SARS-CoV-2) pandemic: A review. *Chaos Solitons Fractals* **2020**, *139*, 110059.
- Shoeibi, A.; Khodatari, M.; Alizadehsani, R.; Ghassemi, N.; Jafari, M.; Moridian, P.; Khadem, A.; Sadeghi, D.; Hussain, S.; Zare, A.; et al. Automated detection and forecasting of COVID-19 using deep learning techniques: A review. *arXiv* **2020**, arXiv:2007.10785.
- Voutilainen, A.; Tolppanen, A.M.; Vehviläinen-Julkunen, K.; Sherwood, P.R. From spatial ecology to spatial epidemiology: Modeling spatial distributions of different cancer types with principal coordinates of neighbor matrices. *Emerg. Themes Epidemiol.* **2014**, *11*, 1–10.
- Desikan, R.; Padmanabhan, P.; Kierzek, A.M.; van der Graaf, P.H. Mechanistic models of COVID-19: Insights into disease progression, vaccines, and therapeutics. *Int. J. Antimicrob. Agents* **2022**, *60*, 106606.
- Friji, H.; Hamadi, R.; Ghazzai, H.; Besbes, H.; Massoud, Y. A generalized mechanistic model for assessing and forecasting the spread of the COVID-19 pandemic. *IEEE Access* **2021**, *9*, 13266–13285.
- Brown, G.D.; Oleson, J.J.; Porter, A.T. An empirically adjusted approach to reproductive number estimation for stochastic compartmental models: A case study of two Ebola outbreaks. *Biometrics* **2016**, *72*, 335–343.
- Shankar, S.; Mohakuda, S.S.; Kumar, A.; Nazneen, P.S.; Yadav, A.K.; Chatterjee, K.; Chatterjee, K. Systematic review of predictive mathematical models of COVID-19 epidemic. *Med. J. Armed Forces India* **2021**, *77*, S385–S392.
- Rahimi, I.; Chen, F.; Gandomi, A.H. A review on COVID-19 forecasting models. *Neural Comput. Appl.* **2021**, 1–11. <https://doi.org/10.1007/s00521-020-05626-8>
- Kong, L.; Duan, M.; Shi, J.; Hong, J.; Chang, Z.; Zhang, Z. Compartmental structures used in modeling COVID-19: A scoping review. *Infect. Diseases Poverty* **2022**, *11*, 72.
- Istaiteh, O.; Owais, T.; Al-Madi, N.; Abu-Soud, S. Machine learning approaches for COVID-19 forecasting. In Proceedings of the 2020 International Conference on Intelligent Data Science Technologies and Applications (IDSTA), Valencia, Spain, 19–22 October 2020; IEEE: Piscataway, NJ, USA, 2020; pp. 50–57.
- Tian, Y.; Luthra, I.; Zhang, X. Forecasting COVID-19 cases using Machine Learning models. *MedRxiv* **2020**. <https://doi.org/10.1101/2020.07.02.20145474>

18. Nikparvar, B.; Rahman, M.; Hatami, F.; Thill, J.C. Spatio-temporal prediction of the COVID-19 pandemic in US counties: Modeling with a deep LSTM neural network. *Sci. Rep.* **2021**, *11*, 1–12.
19. Zeroual, A.; Harrou, F.; Dairi, A.; Sun, Y. Deep learning methods for forecasting COVID-19 time-Series data: A Comparative study. *Chaos Solitons Fractals* **2020**, *140*, 110121.
20. Holmdahl, I.; Buckee, C. Wrong but useful—What COVID-19 epidemiologic models can and cannot tell us. *N. Engl. J. Med.* **2020**, *383*, 303–305.
21. United States Census Bureau. Annual Estimates of the Resident Population for Counties: April 1, 2020 to July 1, 2021 (CO-EST2021-POP). 2021. Available online: <https://www.census.gov/data/tables/time-series/demo/popest/2020s-counties-total.html> (accessed on 11 January 2021).
22. United States Census Bureau. Annual Resident Population Estimates for States and Counties: April 1, 2010 to July 1, 2019; April 1, 2020; and July 1, 2020 (CO-EST2020). 2021. Available online: <https://www.census.gov/programs-surveys/popest/technical-documentation/research/evaluation-estimates/2020-evaluation-estimates/2010s-counties-total.html> (accessed on 11 January 2021).
23. Bureau of Economic Analysis. CAEMP25N Total Full-Time and Part-Time Employment by NAICS Industry 1/Total Employment (Number of Jobs). 2022. Available online: <https://apps.bea.gov/iTable/iTable.cfm?reqid=70&step=1&acrdn=6> (accessed on 10 June 2022).
24. United States Census Bureau. LEHD Origin-Destination Employment Statistics Data (2002–2018). 2018. Available online: <https://lehd.ces.census.gov/data/> (accessed on 5 April 2021).
25. Delmelle, E.C.; Zhou, Y.; Thill, J.C. Densification without growth management? Evidence from local land development and housing trends in Charlotte, North Carolina, USA. *Sustainability* **2014**, *6*, 3975–3990.
26. City of Charlotte. 2030 Transit Corridor System Plan. 2006. Available online: <https://charlottenc.gov/cats/transit-planning/2030-plan/Pages/default.aspx> (accessed on 22 June 2022).
27. Dong, E.; Du, H.; Gardner, L. An interactive web-based dashboard to track COVID-19 in real time. *Lancet Infect. Dis.* **2020**, *20*, 533–534.
28. Centers for Disease Control and Prevention. MMWR Week Fact Sheet. Available online: https://ndc.services.cdc.gov/wp-content/uploads/MMWR_Week_overview.pdf (accessed on 1 April 2022).
29. Pergolizzi JV Jr Magnusson, P.; LeQuang, J.A.; Breve, F.; Paladini, A.; Rekatsina, M.; Yeam, C.T.; Imani, F.; Saltelli, G.; Taylor, R., Jr.; et al. The current clinically relevant findings on COVID-19 pandemic. *Anesthesiol. Pain Med.* **2020**, *10*, e103819.
30. SimplyAnalytics Census 2020 Current Estimates Data. 2020. Available online: <https://simplyanalytics.com/> (accessed on 8 May 2020).
31. Brown, T.S.; Engø-Monsen, K.; Kiang, M.V.; Mahmud, A.S.; Maude, R.J.; Buckee, C.O. The impact of mobility network properties on predicted epidemic dynamics in Dhaka and Bangkok. *Epidemics* **2021**, *35*, 100441.
32. Ramiadantsoa, T.; Metcalf CJ, E.; Raherinandrasana, A.H.; Randrianarisoa, S.; Rice, B.L.; Wesolowski, A.; Randriatsarafara, F.M.; Rasambainarivo, F. Existing human mobility data sources poorly predicted the spatial spread of SARS-CoV-2 in Madagascar. *Epidemics* **2022**, *38*, 100534.
33. Aktay, A.; Bavadekar, S.; Cossoul, G.; Davis, J.; Desfontaines, D.; Fabrikant, A.; Gabrilovich, E.; Gadepalli, K.; Gipson, B.; Guevara, M.; et al. Google COVID-19 Community Mobility Reports: Anonymization process description (version 1.1). *arXiv* **2021**, arXiv:2004.04145.
34. Apple. COVID-19 Mobility Trends Reports. 2021. Available online: <https://COVID19.apple.com/mobility> (accessed on 5 May 2021).
35. SafeGraph. Monthly Patterns. 2021. Available online: <https://docs.safegraph.com/docs/monthly-patterns> (accessed on 5 May 2021).
36. SafeGraph. SafeGraph COVID-19 Data Consortium. 2021. Available online: <https://www.safegraph.com/academics> (accessed on 9 October 2021).
37. Grossman, G.; Kim, S.; Rexer, J.M.; Thirumurthy, H. Political partisanship influences behavioral responses to governors' recommendations for COVID-19 prevention in the United States. *Proc. Natl. Acad. Sci. USA* **2020**, *117*, 24144–24153.
38. North Carolina State Board of Elections. Voting Results. 2020. Available online: <https://er.ncsbe.gov/> (accessed on 17 February 2021).
39. South Carolina Election Commissions. Voting results. 2020. Available online: <https://www.enr-scvotes.org/SC/106502/Web02-state.264691/#/?undefined> (accessed on 21 February 2021).
40. North Carolina Department of Health and Human Services. People Vaccinate Demographics. 2021. Available online: <https://COVID19.ncdhhs.gov/dashboard/data-behind-dashboards> (accessed on 11 December 2021).
41. North Carolina Government. NC COVID-19 Orders and Directives. 2020. Available online: <https://www.nc.gov/COVID-19/COVID-19-orders-directives#executive-orders--2020-executive-orders-issued-by-gov--roy-cooper> (accessed on 21 January 2021).
42. United States Census Bureau. Cartographic Boundary Files—Shapefile. 2018. Available online: <https://www.census.gov/geographies/mapping-files/time-series/geo/carto-boundary-file.html> (accessed on 6 March 2021).

43. Brown, G.D.; Porter, A.T.; Oleson, J.J.; Hinman, J.A. Approximate Bayesian computation for spatial SEIR(S) epidemic models. *Spat. Spatio-Temporal Epidemiol.* **2018**, *24*, 27–37.
44. Toni, T.; Welch, D.; Strelkowa, N.; Ipsen, A.; Stumpf, M.P. Approximate Bayesian computation scheme for parameter inference and model selection in dynamical systems. *J. R. Soc. Interface* **2009**, *6*, 187–202.
45. Ray, E.L.; Wattanachit, N.; Niemi, J.; Kanji, A.H.; House, K.; Cramer, E.Y.; Bracher, J.; Zheng, A.; Yamana, T.K.; Xiong, X.; et al. Ensemble forecasts of coronavirus disease 2019 (COVID-19) in the US. *MedRxiv* **2022**. <https://doi.org/10.1101/2020.08.19.20177493>.
46. Beaumont, M.A. Approximate Bayesian computation in evolution and ecology. *Annu. Rev. Ecol. Evol. Syst.* **2010**, *41*, 379–406.
47. Thill, J.-C. Research on urban and regional systems: Contributions from GIS&T, spatial analysis, and location modeling. In *Innovations in Urban and Regional Systems—Contributions from GIS&T, Spatial Analysis and Location Modeling*; Springer Nature: Cham, Switzerland, 2020; pp. 3–20.
48. Hou, X.; Gao, S.; Li, Q.; Kang, Y.; Chen, N.; Chen, K.; Rao, J.; Ellenberg, J.S.; Patz, J.A. Intracounty modeling of COVID-19 infection with human mobility: Assessing spatial heterogeneity with business traffic, age, and race. *Proc. Natl. Acad. Sci. USA* **2021**, *118*, e2020524118.
49. Lawson, A.B.; Kim, J. Space-time COVID-19 Bayesian SIR modeling in South Carolina. *PLoS ONE* **2021**, *16*, e0242777.
50. Chiang, W.H.; Liu, X.; Mohler, G. Hawkes process modeling of COVID-19 with mobility leading indicators and spatial covariates. *Int. J. Forecast.* **2022**, *38*, 505–520.
51. Kinsey, M.; Tallaksen, K.R.F.O.; Asher, L.; Costello, C.; Kelbaugh, M.; Wilson, S. *Bucky Model: A Spatial SEIR Model for Simulating COVID-19 at the County Level*; The Johns Hopkins University Applied Physics Laboratory LLC: Laurel, MD, USA, 2020. Available online: <https://buckymodel.com/> (accessed on 7 December 2021).
52. Macías, R.Z.; Gutiérrez-Pulido, H.; Arroyo EA, G.; González, A.P. Geographical network model for COVID-19 spread among dynamic epidemic regions. *Math. Biosci. Eng.* **2022**, *19*, 4237–4259.
53. Gopalakrishnan, V.; Pethe, S.; Kefayati, S.; Srinivasan, R.; Hake, P.; Deshpande, A.; Liu, X.; Hoang, E.; Davila, M.; Bianco, S.; et al. Globally local: Hyper-local modeling for accurate forecast of COVID-19. *Epidemics* **2021**, *37*, 100510.
54. Liu, P.; Li, H.X. Global Behavior of a Multi-Group Seir Epidemic Model with Spatial Diffusion in a Heterogeneous Environment. *Int. J. Appl. Math. Comput. Sci.* **2022**, *32*, 271–283.
55. Rajuladevi, R.S.; Marathe, M.; Porebski, P.; Venkataramanan, S. Impact of Seeding and Spatial Heterogeneity on Metapopulation Disease Dynamics. In Proceedings of the epiDAMIK 5.0: The 5th International workshop on Epidemiology meets Data Mining and Knowledge discovery at KDD 2022, Washington, DC, USA, 15 August 2022.
56. Bernardi, M.S.; Sangalli, L.M.; Mazza, G.; Ramsay, J.O. A penalized regression model for spatial functional data with application to the analysis of the production of waste in Venice province. *Stoch. Environ. Res. Risk Assess.* **2017**, *31*, 23–38.
57. Boehm Vock, L.F.; Reich, B.J.; Fuentes, M.; Dominici, F. Spatial variable selection methods for investigating acute health effects of fine particulate matter components. *Biometrics* **2015**, *71*, 167–177.
58. Smith, M.; Fahrmeir, L. Spatial Bayesian variable selection with application to functional magnetic resonance imaging. *J. Am. Stat. Assoc.* **2007**, *102*, 417–431.