Compliance and containment in social distancing: mathematical modeling of COVID-19 across

townships

Abstract: In the early development of COVID-19, large-scale preventive measures, such as border control and air travel restrictions, were implemented to slow international and domestic transmissions. When these measures were in full effect, new cases of infection would be primarily induced by community spread, such as the human interaction within and between neighboring cities and towns, which is generally known as the meso-scale. Existing studies of COVID-19 using mathematical models are unable to accommodate the need for meso-scale modeling, because of the unavailability of COVID-19 data at this scale and the different timings of local intervention policies. In this respect, we propose a meso-scale mathematical model of COVID-19, named the meso-scale Susceptible, Exposed, Infectious, Recovered (MSEIR) model, using town-level infection data in the state of Connecticut. We consider the spatial interaction in terms of the inter-town travel in the model. Based on the developed model, we evaluated how different strengths of social distancing policy enforcement may impact epi curves based on two evaluative metrics: compliance and containment. The developed model and the simulation results help to establish the foundation for community-level assessment and better preparedness for COVID-19.

Keywords: COVID-19; social distancing; epidemic model; spatial interaction; mobility

The year 2020 was deemed to be unprecedented in human history because of the outbreak of the novel and infectious coronavirus (COVID-19). As of early December 2020, the virus has led to over 67 million infections, over 1.5 million deaths, and echoes of economic depression around the globe (Dong, Du, and Gardner 2020). With over 18 million infections and over 300 thousand deaths, the United States has become the largest victim of this public health calamity. The effect of the virus latency, coupled with

the lack of clinical interventions, was further amplified by the understatement of the disease's severity

in its early development in the country. After COVID-19 was declared by the presidential proclamation

as a national emergency on March 1, 2020 (White House 2020), many state and local governments

started to enforce strict preventive measures to mitigate the community spread (Parmet and Sinha

2020).

These preventive measures, known as social distancing (e.g., closure of non-essential

businesses, stay-at-home orders), aim to minimize interpersonal interactions (Gostin and Wiley 2020). It

has been found that these measures have been effective in delaying the spread of the virus by flattening

the epidemic curve (epi curve) through the observation of transmission (Anderson et al. 2020). While

early discussion of social distancing revolved around social impacts such as the economic consequences

(Atkeson 2020; Yang, Zhang, and Chen 2020) and ethical paradoxes (Lewnard and Lo 2020), many recent

studies have integrated social distancing into mathematical epidemic models, attempting to simulate

and predict scenario-based future outbreaks (Chen et al. 2020; Kissler et al. 2020). These models,

however, have been largely focused on the macro-scale using a relatively large geographic unit, such as

country (Gilbert et al. 2020; Kissler et al. 2020), state (Chen et al. 2020), or county (Lai et al. 2020). To

the authors' knowledge, there have been no epidemic models investigating the COVID-19 development

at the meso-scale with a smaller geographic unit, such as town or census tract.

1. The scale issue in modeling COVID-19

The meso-scale, by comparison with the macro-scale (e.g., states) and the micro-scale (e.g.,

individuals), is of critical importance in the effective containment of the epidemic growth. This

significance can be justified by the mechanism of the preventive strategies implemented over the

different phases of epidemic development. In the early development of COVID-19, large-scale

preventive measures, such as border control and air travel restrictions, were implemented to slow

international and domestic transmissions. When these measures were in effect, new cases of infection would be primarily induced by community spread, such as the human interaction within and between neighboring cities, towns, and communities. Existing macro-scale studies using classical epidemic models, notably the Susceptible, Exposed, Infectious, Recovered (SEIR) model, are unable to accommodate the need for meso-scale modeling, because of three existing limitations in COVID-19 research.

First, in the United States, the timing of the COVID-19 outbreak differs by the state as do their regulatory countermeasures, such as the enforcement of stay-at-home orders. It is relatively intractable to model COVID-19 at a macro-scale while considering the heterogeneity in the timing of local policies and the strength of their enforcement. Second, when long-distance travel via flight is restricted, the transmission will be dictated by short-distance travel, such as daily commuting trips via public transit or private automobiles. In this context, individual mobility and the likelihood of travel are largely driven by compliance with social distancing rules. Therefore, modeling COVID-19 at the meso-scale should articulate how social distancing affects people's travel activities or the willingness to travel as parameters to model the process of transmission. This gap has not been fulfilled by the status quo macro-scale models. Third, while COVID-19 data (e.g., infection, death, and recovery) on a daily basis has become largely available in the public sector, data with finer spatial granularities, such as across townships or census tracts, are extremely lacking. These three tiers of research gaps fuel the need to develop a meso-scale epidemic model that simulates past COVID-19 cases while projecting the local, community-level spread in preparing for a pandemic resurgence.

In this paper, we propose a meso-scale epidemic model using town-level COVID-19 infection data in the state of Connecticut. Because the local infection is largely subject to the effects of social distancing, the model development follows two evaluative metrics in social distancing: compliance, which represents the strength of the policy enforcement, and containment, which represents individual

mobility. By incorporating these two metrics into the SEIR model, we have proposed a meso-scale SEIR

model and have performed model fitting and sensitivity analysis under ten different social distancing

scenarios. Using the developed model, we have evaluated how different social distancing strategies (i.e.,

minimal, moderate, and substantial) would shape the COVID-19 epi curve. Most importantly, we have

conducted a field survey to understand people's travel patterns shortly after the outbreak. The survey

results have been incorporated as a mobility parameter into the developed model for simulating

epidemic development. The simulation results could inform both epidemiologists and stakeholders

about the risk of COVID-19 resurgence and has the potential to help identify the etiology of the

transmission across communities.

The paper is organized as follows. Following the background, Section 2 introduces the

methodological development of the model based on the classical SEIR model and the guiding principle

of social distancing. Section 3 applies the new model to the early outbreak in Connecticut, performs the

model fitting, and simulates epi curves at the town level based on different social distancing scenarios

and a proprietary travel survey. Based on the fitted model, the section further simulates the epi curves

and the spatial patterns of town-level infections for the second-wave outbreak. Section 4 discusses the

major findings and insights shed by the modeling results. Lastly, Section 5 concludes the study with long-

term public health impacts.

2. A Meso-scale SEIR model (MSEIR)

2.1 Introduction to SEIR model

Our proposed model stems from the classical SEIR model, a deterministic mathematical model

to simulate epidemiologic dynamics, as shown in Equations (1) through (4). The SEIR model is composed

of four variables: *S* (susceptible population), *E* (exposed population), *I* (infectious population), and *R*

(recovered population). It explicitly quantifies a four-stage cycle of the disease spreading among a

population in terms of differential equations. Each stage is formulated as a derivate of the population

(i.e., S, E, I, R) with respect to time (t), representing the change of the stage-specific population.

$$
\begin{cases}
\dfrac{dS}{dt} = -\dfrac{\beta SI}{N} & (1) \\[2mm]
\dfrac{dE}{dt} = \dfrac{\beta SI}{N} - \sigma E & (2) \\[2mm]
\dfrac{dI}{dt} = \sigma E - \gamma I & (3) \\[2mm]
\dfrac{dR}{dt} = \gamma I & (4)
\end{cases}
$$

In the equations, N denotes the total population ($N = S + E + I + R$); β, σ, and γ are the daily

transmission rate, daily incubation rate, and daily recovery rate, respectively. The basic reproduction

rate can be derived as $R_0 = \beta/\gamma$. This classical SEIR model, along with its many extensions, has been

widely applied to epidemic modeling of COVID-19 (Chen et al. 2020, Kissler et al. 2020, Lai et al. 2020).

2.2 A Conceptual model of social distancing

Adapting the SEIR model to the meso-scale should emphasize the effectiveness of social

distancing in communities. This evaluation follows the Centers for Disease Control and Prevention

(CDC)'s social distancing guidelines for COVID-19 given in three aspects: operations of public facilities,

restrictions on businesses, and restrictions on personal movement (Gostin and Wiley 2020). For the

restrictions on personal movement, the guidelines impose limitations on people's travel and social

behaviors in terms of prohibiting mass gatherings, requiring physical distancing in face-to-face

interaction, and enforcing stay-at-home orders (Gostin and Wiley 2020). The CDC also calls for legal and

community efforts to enhance compliance with these social distancing measures.

Under this guiding principle, we propose a conceptual model in measuring the effectiveness of

social distancing with an emphasis on travel activities as part of the personal movement. The model

comprises two metrics, compliance and containment, that evaluate the effectiveness of the policy

enforcement, as shown in Figure 1. Compliance evaluates the likelihood of the residents not following

the social distancing rules. While there are various ways in which compliance can be articulated, one

variable could be the percentage of residents engaging in travel activities as a surrogate for the metric.

Containment evaluates the level of human mobility, and a variable for the evaluation could be the

maximum distance that people are willing to travel under the social distancing regulation. These two

metrics, resulting from the strengths of the policy enforcement, will likely affect the transmission risk of

the epidemic. This conceptual model featuring the two evaluative metrics is integrated into the classical

SEIR model to develop a meso-scale epidemic model for COVID-19.

[Figure 1 is here]

Figure 1. A conceptual model of the effects of social distancing on travel activities. The solid line

represents the strength of policy enforcement; the dashed line represents the level of transmission risk.

2.3 *Model development*

We modify the SEIR structure with an emphasis on the impacts of travel activities at the meso-

scale, where the study area is a state and the unit of analysis is a town, also known as the county

subdivision in the United States. Recent epidemic models have employed various forms of mobility data,

such as smart-phone heat maps (Lai et al. 2020) and air traffic flow (Gilbert et al. 2020), to estimate

mobility in the SEIR model. Because of the lack of mobility data at the town level, we employ the Huff-

model (Huff 1963) to estimate the potential for travel. The Huff-model is traditionally used for analyzing

the business potential based on the probability of customers' visits to retail and service facilities. It has

also been extended to forecasting the external trips between communities (Anderson 1999; Anderson

2005). The classical Huff-model uses real-world survey data to calibrate the model parameters, including

the attractiveness of facilities and the distance decay (Huff and McCallum 2008). In this paper, we

choose the linear form to estimate the probability of travel between two towns, as shown in Equation

(5). This linear form is seen as a better form to forecast inter-city trips and has been corroborated with

field data (Anderson 1999).

$$T_{ij} = \frac{\frac{N_i}{D_{ij}}}{\Sigma_j \frac{N_i}{D_{ij}}} \qquad (5)$$

where T_{ij} is the probability that a person traveling from town i to town j, D_{ij} is the distance between i and

j, and N_i is the total population of town i.

We have further added to the Huff model two other parameters: a compliance parameter C_i,

meaning the percentage of the population of town i engaging in inter-town trips, and a containment

parameter D_0, meaning the maximum distance people are willing to travel under influences of social

distancing. These two parameters extend the Huff-model to estimate M_{ij}, the total population traveling

from town i to town j, as shown in Equation (6).

$$M_{ij} = N_i C_i T_{ij} = N_i C_i \frac{\frac{N_i}{D_{ij}}}{\Sigma_j \frac{N_i}{D_{ij}}} \qquad (6)$$

$$\text{s.t.} \qquad D_{ij} \leq D_0$$

Equation (6) is a Huff-based trip distribution model where the compliance parameter C_i and the

containment variable D_0 can be determined by different social distancing scenarios. To further

incorporate the trip distribution model to the SEIR model, we have made several necessary

assumptions: (1) both the susceptible and exposed populations conform to the mobility rule in Equation

(6); (2) the infectious and recovered populations are isolated so that they cannot travel to other towns;

(3) the susceptible population traveling to other towns can be affected by the infectious population in

both their origin town and destination town; (4) the total population and the daily traveling population

of a town are stable during the modeling period; (5) daily travelers return to their origin town by the end

of the day; (6) the transmission rate gradually decreases due to non-pharmaceutical interventions

during the social distancing period (Lai et al. 2020). Based on these assumptions, we have developed the

meso-scale SEIR model (MSEIR), simulating the daily dynamics of susceptible (S_i), exposed (E_i), infectious

(I_i), and recovered populations (R_i) of the ith town, as shown in Equations (7) through (11).

$$
\begin{cases}
\dfrac{dS_i}{dt} = -\dfrac{\beta S_i I_i}{N_i} - \sum_j \dfrac{\beta M_{ij} S_i I_j}{(N_i - I_i - R_i)N_j} & (7) \\[2ex]
\dfrac{dE_i}{dt} = \dfrac{\beta S_i I_i}{N_i} - \sigma E_i + \sum_j \dfrac{\beta M_{ij} S_i I_j}{(N_i - I_i - R_i)N_j} & (8) \\[2ex]
\dfrac{dI_i}{dt} = \sigma E_i - \gamma I_i & (9) \\[2ex]
\dfrac{dR_i}{dt} = \gamma I_i & (10) \\[2ex]
\dfrac{d\beta}{dt} = -a\beta & (11)
\end{cases}
$$

where

E_i: exposed population of town i;

I_i: infectious population of town i;

M_{ij}: population (susceptible or exposed) traveling from town i to town j. The parameter is derived from

Equation (6).

N_i: total population of town i ($N_i = S_i + E_i + I_i + R_i$);

R_i: recovered population of town i (including hospitalized, self-recovered, and death);

S_i: susceptible population of town i;

t: time (daily);

a: daily change rate of the transmission rate ($0 < a < 1$);

β: transmission rate (conversion from the susceptible population to exposed population);

γ: recovery rate;

σ: daily incubation rate (reciprocal of the incubation period);

In the MSEIR model, Equations (7) through (10) are a series of differential equations indicating the daily variation of the susceptible, exposed, infectious, and recovered population at each transmission stage. Equation (11) is the daily change of the transmission rate.

2.4 *Model initialization and parameter estimation*

Five state variables must be initiated prior to the model implementation: (1) the initial transmission rate β_0 (estimated from our case study, see below); (2) the initial exposed population E_{i0} (estimated from our case study, see below); (3) the initial infectious population I_{i0}, which equals to the cases of infection on March 23, 2020, the start date of the data; (4) the initial susceptible population S_{i0}, which can be derived as $S_{i0} = N_i - E_{i0}$; and (5) the initial recovered population $R_0 = 0$ under the assumption that no individuals are cured, hospitalized, or had died at the initial stage.

σ and γ can be derived from historical data. σ is the daily incubation rate as the reciprocal of the incubation period. γ is the recovery rate, indicating the rate of reduction in the infectious population due to hospitalization, self-recovery, and death. This parameter assumes that once an infectious individual is hospitalized, self-recovered, or died, the person will be isolated from the transmission cycle. According to a recent study among the first 425 diagnosed patients (Li et al. 2020), the mean incubation period of COVID-19 was 5.2 days (at a 95% confidence interval [CI], 4.1 to 7.0 days), and the mean time duration from the illness onset to hospital admission was 9.1 days (95% CI, 8.6 to 9.7 days). Thus, we assumed that our study had a similar incubation period and duration from illness onset to the first medical visit. In our model, $\sigma = 1/5.2$ and $\gamma = 1/9.1$.

For parameters β_0 and a, we estimated their optimal values for each town using the daily cumulative cases of infection derived from the Connecticut Department of Public Health (CTDPH 2020).

We assumed that β_0 and a were constant throughout the study area and under different social

distancing scenarios. The exposed population at the initial stage E_{i0} was an unknown parameter that

varied by the town. Because of this uncertainty, we treated E_{i0} as another parameter to be estimated.

Thus, each town i had an independent E_{i0} given the difference in the onset of the outbreak, population,

and other factors dictating the early exposed population. The Nelder-Mead algorithm (Nelder and Mead

1965) was employed to estimate parameters by minimizing the negative normal log-likelihood between

the simulated and the confirmed daily cumulative cases.

3. Case study

3.1 Study area and data

Our first round of simulations was focused on the early outbreak in the state of Connecticut with

the unit of analysis being a town. Located in the New England region, Connecticut is the third smallest

state by area in the United States with 169 towns and a total population of 3.5 million (Figure 2a). On

March 8, the first COVID-19 case was reported in Wilton, a town neighboring New York (The New York

Times 2020). Because of the geographical proximity to New York City, the epicenter of the national

outbreak, the state experienced an exponential rise of infections in the early outbreak. As of May 11,

2020, the total confirmed cases of infection were over 34,000, and the total deaths were over 3,000

(CTDPH 2020). Figure 2 shows the population density as well as the COVID-19 infection rate in the early

outbreak.

[Figure 2 is here]

Figure 2. Connecticut towns with (a) population density and (b) COVID-19 infection rate as of May 11.

Towns further discussed in the article are labeled.

Modeling COVID-19 at the meso-scale is inseparable from state policy related to social

distancing. In Connecticut, the initial social distancing rules were implemented on March 23 by the

governor's executive order "Stay Safe, Stay Home," requiring the closure of non-essential businesses

and some non-profit organizations (Ct.gov 2020a). This policy was relieved by reopening certain non-

essential businesses effective on May 20, marking the inception of the Phase 1 reopening plan (Ct.gov

2020b). In this context, we retrieved the daily town-level COVID-19 infection data before the Phase 1

period, specifically the first 50 days since the state's social distancing rules were in effect (i.e., March 23

through May 11). The COVID-19 dataset was solicited from the state government's daily publications

(CTDPH 2020).

3.2 Social distancing scenarios

We have designed three compliance levels and three containment levels to estimate M_{ij} in the

MSEIR model, forming a total of nine models representing different degrees of social distancing policy

enforcement, as shown in Table 1. In this framework, Model 1 represents the substantial enforcement

with only 10% of the population taking inter-town trips and a maximum travel distance of 20 miles (1

mile = 1.609 km); Model 9 represents the minimum enforcement with 50% of the population taking

inter-town trips and a maximum travel distance of 140 miles. The 140-mile threshold is the road

network distance between the two most remote towns in Connecticut (i.e., Thompson and Greenwich).

Table 1. Social distancing scenarios based on different compliance and containment levels.

Compliance level (C_i)	Containment level (D_0)		
	20 miles	60 miles	140 miles
10%	Model 1 (Substantial)	Model 2	Model 3
30%	Model 4	Model 5 (Moderate)	Model 6
50%	Model 7	Model 8	Model 9 (Minimum)

These social distancing scenarios were incorporated into Equation (6) by Python scripting to

derive $M_{ij.}$ in the implementation, P_i was derived from the 2018 census data; D_{ij} was derived as the road

network distance between the geographic centers of towns using the Network Analysis module in ESRI

ArcMap 10.7 (i.e., OD cost matrix) in a refined road network. The results of M_{ij} were imported to the

MSEIR model (scripted in Python) for fitting the dynamics of COVID-19 infection in the 50-day early

outbreak period. We then employed the fitted models for simulating the trends.

3.3 Results

Using the town-level data, we implemented the MSEIR model under all social distancing

scenarios in Table 1. In addition to the model fitting, we extended the epi curves of the cumulative cases

of infection (I_C) to the epidemic development with an end date of July 12. We also established the

baseline scenario in the SEIR model where there is no travel or interaction between towns (C_i = 0%, D_0 =

0 miles).

For the sake of clarity, we selected six scenarios for comparison and evaluation: the SEIR model for no interaction and Models 1, 3, 5, 7, and 9 with varying degrees of social distancing. Figure 3 shows the simulated epi curves of the entire state. As shown in the figure, the SEIR model has considerably underestimated the epi curve, with the value of I_c converging to 26,000. When the social distancing scenarios are introduced, the epi curves start to align with the confirmed cases, with the minimum social distancing scenario (i.e., Model 9) yielding the steepest curve. It is worth noting that the observed cases as of July 12 were 47,510, which is within the range of Model 5 (45,752) and Model 9 (48,105).

[Figure 3 is here]

Figure 3. Simulated cumulative cases of infection (I_c) for the state by the SEIR model and the MSEIR model (Models 1, 3, 5, 7, and 9) based on confirmed cases until May 11.

[Figure 4 is here]

Figure 4. Simulated cumulative cases of infection (I_c) by the SEIR model and the MSEIR models for (a) Hampton, (b) Franklin, (c) Wilton, (d) Mansfield, (e) New Britain, and (f) Hartford. Simulations were based on confirmed cases until May 11.

Figure 4 shows the epi curves for six selected towns, where their geographical locations are

given in Figure 2. The towns were selected based on the urban-rural classification given by USCB (2010):

Hampton (population: 1853) and Franklin (population: 1933) are rural areas; Wilton (population: 18,397)

and Mansfield (population: 25,817) are urban clusters, where Mansfield comprises mainly university

employees and students; Hartford (population: 122,587), the state capital, and its satellite city New

Britain (population: 72,453) are urbanized areas. It can be seen from the results that the SEIR model

with no spatial interaction generates the flattest curve for all given towns except for Hampton (Figure

4a). The results under different social distancing scenarios are mixed. Specifically, for Hampton, the

observed cases fall below all estimations; for Franklin, the observed cases are between the SEIR model

and Model 1; for Wilton and Mansfield, the observed cases align with the SEIR model; for New Britain

and Hartford, all simulation results are under-estimated. We feel these uncertainties could be explained

by the discrepancies between the simulated travel flow and the real-world mobility patterns, where

travel activities in urbanized areas are more intense than what the model simulates.

3.4 Evaluation of model fitting

To evaluate the performance of the MSEIR model and its variations, we compared Models 1-9

with the original SEIR model. It is worth noting that a fundamental statistical bias, called the edge effect,

exists in the modeling results—for analyses conducted within a finite geographic region, spatial

interactions with entities beyond the region are overlooked (Chen 2017; Vidal Rodeiro and Lawson

2005). In our case, the virus transmission could be introduced by interstate travel, where towns on the

borders were largely affected. To alleviate the edge effect, we excluded the border towns (n = 29) and

focused on the rest of the towns (n = 140) in the evaluation.

We employed two statistical metrics for each of the ten models: r^2 and *root-mean-square error*

(*RMSE*). Specifically, for each model, we calculated r^2 and *RMSE* for each of the 140 towns by comparing

the simulation results with the confirmed cumulative cases in the 50-day period. r^2 assesses whether the

MSEIR model captures the trend of the historical data, while *RMSE* quantifies the absolute difference

between the model output and the observation. Then, for each model, we performed the evaluations

across all 140 towns and derived the average of the 140 values of r^2 and *RMSE* (Table 2). As these two

metrics cannot be intuitively interpreted, we introduced another metric, named the improvement ratio,

to reveal the improvement of the MSEIR models over the SEIR model. The ratio is defined as the

percentage of the towns where the RMSE in the MSEIR model is less than that in the SEIR model, as

shown in Equation (12) and Table 2. We also employed the paired t-test to evaluate the difference

between the RMSE values of the SEIR model and the RMSE values of each MSEIR model, as shown in

Table 2.

$$Improvement\ ratio = \frac{Town\ count\ (MSEIR\ RMSE\ <\ SEIR\ RMSE)}{Total\ town\ count} * 100\% \qquad (12)$$

Overall, all models have achieved satisfactory levels of fitting with the historical data in terms of

trend fitting ($r^2 > 0.9$). However, in terms of fitting the absolute values by the average *RMSE*, the SEIR

model is not satisfactory, where the average *RMSE* (72.595) is the largest, meaning that its simulated

cases largely deviate from the observation. We also identified that Model 7 performs better than other

models in terms of a relatively high r^2 (0.918), the least average *RMSE* (46.166), the best improvement

ratio (69%) over the SEIR model, and one of the highest significance levels in the paired t-test (p-value <

0.001). We thus choose Model 7 for simulating the second-wave outbreak (see Section 3.6).

To further explore the applicability of the MSEIR model, we divided the towns into three

categories based on the United States Census Bureau (USCB)'s urban-rural classification (USCB 2010):

urbanized areas (UAs), urban clusters (UCs), and rural areas (RAs). Then, we evaluated the model fitting

for towns under each category, as shown in Table 3. We identified that for UAs and UCs, the models

achieve satisfactory levels of fitting ($r^2 > 0.9$), and they have the best performance for UAs in terms of

relatively high improvement ratios; for RAs, the observation cannot be fitted well ($r^2 < 0.7$). We feel this

discrepancy could be introduced by the relatively limited number of observed cases in small towns,

giving rise to the uncertainty of outbreaks and the difficulty for model simulation. Thus, in general, the

MSEIR model is best suited for towns exceeding a certain size and must be adjusted with real-world trip

data for applications to small towns or rural areas.

Table 2. Model fitting with confirmed cases for all towns (n = 140).

Model	Average r^2	Average $RMSE$	Improvement ratio	Paired t-test t-value
SEIR	0.924	72.595	N/A	N/A
1	0.912	66.987	62.1%	0.990
2	0.913	68.315	53.6%	0.698
3	0.915	59.076	62.1%	2.530*
4	0.914	48.062	67.1%	4.032***
5	0.907	58.816	57.1%	1.848
6	0.912	47.604	57.9%	4.360***
7	0.918	46.166	69.3%	4.296***
8	0.918	56.222	39.3%	1.970
9	0.915	55.335	41.4%	2.489*

*p-value < 0.05 ***p-value < 0.001

Table 3. Model fitting with confirmed cases for towns by category.

Model	UAs (*n* = 15)				UCs (*n* = 113)				RAs (*n* = 12)			
	Average r^2	Average RMSE	Improvement ratio	Paired t-test t-value	Average r^2	Average RMSE	Improvement ratio	Paired t-test t-value	Average r^2	Average RMSE	Improvement ratio	Paired t-test t-value
SEIR	0.987	319.891	N/A	N/A	0.949	47.153	N/A	N/A	0.603	3.057	N/A	N/A
1	0.979	261.351	60.0%	1.747	0.936	48.080	62.8%	-0.178	0.613	2.076	58.3%	1.687
2	0.983	286.676	66.7%	0.684	0.937	46.015	52.2%	0.273	0.608	5.352	50.0%	-1.100
3	0.984	250.103	66.7%	1.881	0.937	39.645	63.7%	1.831	0.621	3.268	41.7%	-0.230
4	0.982	186.627	80.0%	3.337**	0.937	34.575	66.4%	3.287**	0.609	1.859	58.3%	1.934
5	0.978	251.092	80.0%	1.157	0.928	39.247	55.8%	1.658	0.619	2.748	41.7%	0.449
6	0.978	164.976	100.0%	6.037***	0.935	36.652	54.9%	2.543*	0.613	4.015	33.3%	-1.098
7	0.986	176.308	93.3%	3.599**	0.942	33.568	68.1%	3.706***	0.613	2.112	50.0%	1.231
8	0.988	135.591	93.3%	4.264***	0.942	50.886	36.3%	-0.607	0.603	7.265	0.0%	-5.98***
9	0.984	157.769	93.3%	5.277***	0.939	46.981	36.3%	0.031	0.612	5.954	25.0%	-2.861*

Note: The categories are given by USCB (2010) based on population: urbanized areas (UAs) of 50,000 or more people, urban clusters (UCs) between 2,500 and 50,000 people, and rural areas (RAs) of less than 2,500 people. It should be noted that USCB uses a different areal unit for urban-rural classification and is not based on the county subdivision (i.e., town). Thus, the category in this analysis does not suggest the actual urban-rural status.

*p-value < 0.05 **p-value < 0.005 ***p-value < 0.001

3.5 *Calibrating MSEIR model with travel survey*

To further explore these model discrepancies and better estimate the mobility parameter, specifically the M_{ij} in the MSEIR model, we launched a field survey to investigate the real-world travel activities in the state. The survey was distributed on Amazon's MTurk and was reposted weekly, lasting for a total of 74 days (i.e., June 2 through August 14). Only Connecticut residents were eligible to take the survey. The survey questions included the participants' residential town, if they have traveled to another town in the last week, towns to which they have traveled, and the number of trips for each town they have traveled to. Responses solicited from the survey were used to estimate the frequency of daily inter-town trips or the real-world M_{ij} (M_{ij}^R) in the MSEIR model. The estimation is based on the following assumptions: (1) participants' travel activities represented the mobility pattern of their residential town; (2) the total population and the daily traveling population of a town were stable during the survey period; (3) trips across different weeks were independent. Based on these assumptions, we propose Equation (13) to estimate M_{ij}^R.

$$M_{ij}^R = \frac{m_{ij}N_i}{dP_i} \qquad (13)$$

where

d: number of days during the survey period;

m_{ij}: total number of trips from town i to town j taken by survey participants during the survey period;

N_i: total population of town i;

P_i: total number of survey participants residing in town i;

After eliminating incomplete answers, we received 1225 valid responses, accounting for a total of 4000 completed trips. The survey data were employed to estimate M_{ij}^R based on Equation (13), which was further incorporated into the MSEIR for model calibration and comparison.

3.6 Second-wave simulations

On June 17, 2020, the state government started to enforce the Phase 2 reopening plan, allowing the limited operation of certain non-essential businesses (ct.gov 2020b). Because this change in social distancing policy could affect the mobility pattern, we initiated another round of simulations with confirmed cases drawn from the Phase 2 period. We chose Model 7 (i.e., model with the best performance) and Model R (i.e., model with real-world trip data) for simulating the second wave of the outbreak. Additionally, responses to our travel survey were mostly collected during this period, allowing for the justification of the mobility parameter or M_{ij}^R in Model R. Thus, we refitted the model parameters using the confirmed cases from June 17 through September 7 and then changed the model initialization parameters based on June 17. The simulation results for the period of September 8 through December 31 are given in Figures 5 through 7.

Figure 5 shows the epi curves simulated by Model 7 and Model R for the state. The epi curve of Model 7 well aligns with the trend, and it indicates that the cumulative cases of infection would be 73,866 by end of 2020 if the same level of policy enforcement remains in effect. However, the epi curve of Model R is relatively flat and underestimates the trend. This underestimation could be attributed to the relatively small sample in the travel survey. Figure 6 shows the simulated epi curves for the six selected towns. It could be seen from the two figures that comparing with the observed cases, the MSEIR model considerably underperforms on both the state level and the town level. The underperformance is likely due to the drastic change of social distancing policy, where the state entered the Phase 3 reopening on October 8 and then retreated to Phase 2.1 on November 6 because of a surge in infections (ct.gov 2020c). Therefore, the MSEIR model is unable to characterize the epi curve if major changes in social distancing policy occur.

In addition, Figure 7 visualizes the projected spatial patterns of the infection under the premise that the same level of social distancing is enforced. The figure shows the simulated infection rate and

the simulated rate of increase (based on September 7) across all towns by Model 7 as of December 31.

We excluded the border towns in the figure, since their simulation results are subject to the edge effect.

Towns with the top-10 highest rates in the two maps are labeled. Figure 7a shows that towns along the

southwestern coast, towns near Hartford, and the town of Montville have alarming rates of outbreaks.

Figure 7b shows several high-risk towns in the eastern part of the state—especially, Ledyard and Lisbon.

Health policy intervention, such as increasing the coverage of testing, should thus be shifted towards

these high-risk towns.

[Figure 5 is here]

Figure 5. Simulated cumulative cases of infection (I_c) for the state by the MSEIR model (Models R and 7)

based on confirmed cases until September 7.

[Figure 6 is here]

Figure 6. Simulated cumulative cases of infection (I_c) by the SEIR model and the MSEIR models for (a)

Hampton, (b) Franklin, (c) Wilton, (d) Mansfield, (e) New Britain, and (f) Hartford. Simulations were

based on confirmed cases until September 7.

[Figure 7 is here]

Figure 7. (1) Simulated infection rate and (2) simulated rate of increase by Model 7 as of December 31.

Towns with the top-10 highest rates are labeled.

4. Discussion

The proposed MSEIR model applied to meso-scale COVID-19 simulations is among the first to

evaluate the development of the pandemic using an administrative unit smaller than a county. By

downscaling the analysis to the town level and realizing the model under different social distancing

scenarios, the study sheds important insights into COVID-19 studies.

First, meso-scale analysis is of critical importance for revealing the epidemic development after

the initial outbreak. When social distancing orders were placed to curb the early infection, needs for

domestic flights or interstate travel were largely suppressed (Gao et al. 2020). While long-distance travel

was restricted, new cases of infection were primarily caused by local spread through short-distance

travel. Thus, efforts and policies to contain the COVID-19 development would be most effective by

curbing inter-and intra-town travel activities. Incorporating the interaction across townships and

deriving their epi curves can help the municipality to leverage resources for preparing for rising

contingencies, such as the resurgence of outbreaks. The classical SEIR model and its many extensions,

however, lack the capability of simulating the epidemic spread at the meso-scale. This increased spatial

granularity to model COVID-19 is the major contribution of this work.

Second, the proposed social distancing framework including the compliance and containment

provides quantifiable metrics for COVID-19 studies that attempt to evaluate the effects of social

distancing. Since the pandemic is growing at an alarming rate worldwide, existing studies have largely

emphasized the timing (Chinazzi et al. 2020), economic impacts (Atkeson 2020), and the ethical issues

(Lewnard and Lo 2020) of social distancing, while the effects on human mobility at the community scale

are not well scrutinized. This gap has likely resulted from the lack of granular mobility data (especially

the origin-destination trip data), coupled with the sensitivity of health data collection at a refined scale

(e.g., the Health Insurance Portability and Accountability Act). The simulation of the travel activities

using the Huff model could help to estimate the regional human movement pattern, and the model

calibration using the survey method will increase the rigor of the estimation.

Third, for the early outbreak, we have found that there are discrepancies in the modeling results

across different towns. While the SEIR model has a considerably poor fitting, it tends to align with towns

with a sparse population. This alignment is very likely due to the relatively low level of spatial interaction in terms of inter-town travel in these towns. The proposed MSEIR model works best for towns exceeding a certain size (e.g., UCs and UAs). This result justifies that although towns in a state are subject to the same timing of social distancing orders, the actual policy effects on residents' mobility, and consequently, on curbing the epi curves of the pandemic are seemingly different. Due to this spatial heterogeneity in the mobility pattern, which is internally driven by socioeconomic inequities (Bonaccorsi et al. 2020), it is impossible to establish a one-size-fits-all model for COVID-19 analysis for every town in a state. Therefore, we have two recommendations for improving and better applying the MSEIR model: first, a travel survey or equivalent trip generation data (e.g., SafeGraph mobility data) is a necessity to derive the actual mobility pattern at the meso-scale and estimate the MSEIR model parameters; second, if such real-world mobility data are not attainable, we suggest that local stakeholders employing the model should adopt and prepare for the worst-case scenario (e.g., Model 9) and target towns that may experience the most rapid epidemic growth under all scenarios. This elevated caution can guide the leverage of public health resources towards the most severe pandemic situation.

Lastly, if any major changes in social distancing policy occur, the proposed MSEIR will be less effective. With the rapidly evolving pandemic and varying degrees of social distancing and reopening policies, projecting the epi curves in the months to come is beyond the model's capacity. For example, in Connecticut, the total cases of infection were over 135,000 as of December 7, which outnumbered any simulation results. This steep rise was likely due to the implementation of subsequent reopening plans (Ct.gov 2020c), which largely increased the capacities of non-essential businesses and, therefore, the likelihood of transmission. Moreover, follow-up events, including the presidential election and national holidays could have contributed to the spike in infections (Cotti et al. 2020). To minimize these uncertainties, the model implementation should be focused on simulating cases in a short future period (e.g., one month) and should be conducted on a weekly basis to reflect the changes in public policies.

5. Conclusion

The COVID-19 pandemic has posed an unprecedented challenge to the global economy and the healthcare system. While modeling COVID-19 by simulating the epi curve has become a growing practice across all disciplines, existing models have not been able to examine the issue at the meso-scale, using a small unit of analysis such as town or census tract, nor have they quantified how social distancing may curb the transmission at this scale. The proposed MSEIR model introduces the effects of spatial interaction between towns on the epidemic development. The scenario-based analysis could help policy stakeholders to understand how the compliance with and the containment by social distancing rules regulate people's travel activities and can help predict how different degrees of policy enforcement would shape the epi curve. These modeling results have the potential to assist stakeholders with strategical decisions about the timing and expected outcomes of relieving social distancing requirements. Eventually, we believe the developed MSEIR model will establish the scientific foundation for community-level assessment and better preparedness for COVID-19.

Acknowledgment

Disclosure statement

No potential conflict of interest was reported by the authors. The travel survey in the study was approved by the Institutional Review Board (IRB) at University of Connecticut under Protocol X20-0096.

Data and Codes Availability Statement

The dataset and codes for the MSEIR model are available in Github [https://github.com/peterbest52/mseir]. The COVID-19 dataset was derived from the public domain: CT Coronavirus daily report [https://portal.ct.gov/coronavirus]. The travel survey data were collected by the research team.

References:

Anderson, M. D. 1999. Evaluation of models to forecast external-external trip percentages. *Journal of Urban Planning and Development*, 125(3), 110–120.

Anderson, M. D. 2005. Spatial economic model for forecasting the percentage splits of external trips on highways approaching small communities. *Transportation Research Record*, *1931*(1), 68–73.

Anderson, R. M., Heesterbeek, H., Klinkenberg, D., & Hollingsworth, T. D. 2020. How will country-based mitigation measures influence the course of the COVID-19 epidemic?. *The Lancet*, *395*(10228), 931–934.

Atkeson, A. 2020. What will be the economic impact of COVID-19 in the US? Rough estimates of disease scenarios (No. w26867). *National Bureau of Economic Research*.

Bonaccorsi, G., Pierri, F., Cinelli, M., Porcelli, F., Galeazzi, A., Flori, A., ... & Pammolli, F. 2020. Evidence of economic segregation from mobility lockdown during COVID-19 epidemic. *Available at SSRN 3573609*.

Chen, S., Li, Q., Gao, S., Kang, Y., & Shi, X. 2020. Mitigating COVID-19 outbreak via high testing capacity and strong transmission-intervention in the United States. *medRxiv*.

Chen, X. 2017. Take the edge off: A hybrid geographic food access measure. *Applied Geography*, 87, 149–159.

Chinazzi, M., Davis, J. T., Ajelli, M., Gioannini, C., Litvinova, M., Merler, S., ... & Viboud, C. 2020. The effect of travel restrictions on the spread of the 2019 novel coronavirus (COVID-19) outbreak. *Science*, *368*(6489), 395□–400.

Connecticut Department of Public Health (CTDPH). 2020. Test data in Connecticut [online]. Available from: https://portal.ct.gov/Coronavirus [Accessed 10 December 2020].

Cotti, C. D., Engelhardt, B., Foster, J., Nesson, E. T., & Niekamp, P. S. (2020). The relationship between in-person voting and COVID-19: Evidence from the Wisconsin primary. *NBER Working Paper*, 27187.

Ct.gov. 2020a. Governor Lamont releases guidance to businesses on order asking Connecticut to 'Stay Safe, Stay Home" [online]. Available from: https://portal.ct.gov/Office-of-the-Governor/News/Press-Releases/2020/03-2020/Governor-Lamont-Releases-Guidance-to-Businesses-on-Order-Asking-Connecticut-to-Stay-Safe-Stay-Home [Accessed 10 December 2020].

Ct.gov. 2020b. Governor Lamont releases rules for businesses under first phase of Connecticut's reopening plans amid COVID-19 [online]. Available from: https://portal.ct.gov/Office-of-the-Governor/News/Press-Releases/2020/05-2020/Governor-Lamont-Releases-Rules-for-Businesses-Under-First-Phase-of-Reopening-Plans [Accessed 21 December 2020].

Ct.gov. 2020c. Sector rules and certification for reopen [online]. Available from: https://portal.ct.gov/DECD/Content/Coronavirus-Business-Recovery/Sector-Rules-and-Certification-for-Reopen [Accessed 21 December 2020].

Dong, E., Du, H., & Gardner, L. 2020. An interactive web-based dashboard to track COVID-19 in real time. *The Lancet Infectious Diseases*.

Gao, S., Rao, J., Kang, Y., Liang, Y., & Kruse, J. 2020. Mapping county-level mobility pattern changes in the United States in response to COVID-19. *SIGSPATIAL Special*, 12(1), 16–26.

Gilbert, M., Pullano, G., Pinotti, F., Valdano, E., Poletto, C., Boëlle, P. Y., ... & Gutierrez, B. 2020. Preparedness and vulnerability of African countries against importations of COVID-19: A modelling study. *The Lancet*, *395*(10227), 871–877.

Gostin, L. O., & Wiley, L. F. 2020. Governmental public health powers during the COVID-19 pandemic: stay-at-home orders, business closures, and travel restrictions. *JAMA*, *323*(21), 2137–2138.

Huff, D. L. 1963. A probabilistic analysis of shopping center trade areas. *Land Economics*, *39*(1), 81–90.

Huff, D., & McCallum, B. M. 2008. Calibrating the huff model using ArcGIS business analyst. *ESRI White Paper*.

Kissler, S. M., Tedijanto, C., Lipsitch, M., & Grad, Y. 2020. Social distancing strategies for curbing the COVID-19 epidemic [online]. *medRxiv*. Available from: https://www.medrxiv.org/content/10.1101/2020.03.22.20041079v1 [Accessed 10 December 2020].

Lai, S., Ruktanonchai, N. W., Zhou, L., Prosper, O., Luo, W., Floyd, J. R., ... & Yu, H. 2020. Effect of non-pharmaceutical interventions to contain COVID-19 in China [online]. Available from: https://dash.harvard.edu/handle/1/42661263 [Accessed 10 December 2020].

Lewnard, J. A., & Lo, N. C. 2020. Scientific and ethical basis for social-distancing interventions against COVID-19. *The Lancet Infectious Diseases*, 20(6), 631–633.

Li, R., Pei, S., Chen, B., Song, Y., Zhang, T., Yang, W., & Shaman, J. 2020. Substantial undocumented infection facilitates the rapid dissemination of novel coronavirus (SARS-CoV-2). *Science*, *368*(6490), 489–493.

Nelder, J. A., & Mead, R. 1965. A simplex method for function minimization. *The Computer Journal*, 7(4), 308–313.

Parmet, W. E., & Sinha, M. S. 2020. Covid-19—the law and limits of quarantine. *New England Journal of Medicine*, *382*(15), e28.

The New York Times. 2020. Coronavirus in Connecticut: governor announces first case of resident [online]. Available from: https://www.nytimes.com/2020/03/08/nyregion/coronavirus-connecticut.html [Accessed 10 December 2020].

US Census Bureau (USCB). 2010. Urban and rural [online]. Available from: https://www.census.gov/programs-surveys/geography/guidance/geo-areas/urban-rural.html [Accessed 10 December 2020].

Vidal Rodeiro, C. L., & Lawson, A. B. 2005. An evaluation of the edge effects in disease map modelling. Computational Statistics and Data Analysis, 49(1), 45–62.

White House. 2020. Proclamation on declaring a national emergency concerning the Novel Coronavirus Disease (COVID-19) outbreak [online]. Available from https://www.whitehouse.gov/presidential-actions/proclamation-declaring-national-emergency-concerning-novel-coronavirus-disease-covid-19-outbreak/ [Accessed 10 December 2020].

Yang, Y., Zhang, H., & Chen, X. (2020). Coronavirus pandemic and tourism: Dynamic stochastic general equilibrium modeling of infectious disease outbreak. *Annals of Tourism Research*, 83, 102913.

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Table 1. Social distancing scenarios based on different compliance and containment levels.

Compliance level (C_i)	Containment level (D_0)		
	20 miles	60 miles	140 miles
10%	Model 1 (Substantial)	Model 2	Model 3
30%	Model 4	Model 5 (Moderate)	Model 6
50%	Model 7	Model 8	Model 9 (Minimum)

Table 2. Model fitting with confirmed cases for all towns (n = 140).

Model	Average r^2	Average $RMSE$	Improvement ratio	Paired t-test t-value
SEIR	0.924	72.595	N/A	N/A
1	0.912	66.987	62.1%	0.990
2	0.913	68.315	53.6%	0.698
3	0.915	59.076	62.1%	2.530*
4	0.914	48.062	67.1%	4.032***
5	0.907	58.816	57.1%	1.848
6	0.912	47.604	57.9%	4.360***
7	0.918	46.166	69.3%	4.296***
8	0.918	56.222	39.3%	1.970
9	0.915	55.335	41.4%	2.489*

*p-value < 0.05 ***p-value < 0.001

Table 3. Model fitting with confirmed cases for towns by category.

Model	UAs (n = 15)				UCs (n = 113)				RAs (n = 12)			
	Average r^2	Average RMSE	Improvement ratio	Paired t-test t-value	Average r^2	Average RMSE	Improvement ratio	Paired t-test t-value	Average r^2	Average RMSE	Improvement ratio	Paired t-test t-value
SEIR	0.987	319.891	N/A	N/A	0.949	47.153	N/A	N/A	0.603	3.057	N/A	N/A
1	0.979	261.351	60.0%	1.747	0.936	48.080	62.8%	-0.178	0.613	2.076	58.3%	1.687
2	0.983	286.676	66.7%	0.684	0.937	46.015	52.2%	0.273	0.608	5.352	50.0%	-1.100
3	0.984	250.103	66.7%	1.881	0.937	39.645	63.7%	1.831	0.621	3.268	41.7%	-0.230
4	0.982	186.627	80.0%	3.337**	0.937	34.575	66.4%	3.287**	0.609	1.859	58.3%	1.934
5	0.978	251.092	80.0%	1.157	0.928	39.247	55.8%	1.658	0.619	2.748	41.7%	0.449
6	0.978	164.976	100.0%	6.037***	0.935	36.652	54.9%	2.543*	0.613	4.015	33.3%	-1.098
7	0.986	176.308	93.3%	3.599**	0.942	33.568	68.1%	3.706***	0.613	2.112	50.0%	1.231
8	0.988	135.591	93.3%	4.264***	0.942	50.886	36.3%	-0.607	0.603	7.265	0.0%	-5.98***
9	0.984	157.769	93.3%	5.277***	0.939	46.981	36.3%	0.031	0.612	5.954	25.0%	-2.861*

Note: The categories are given by USCB (2010) based on population: urbanized areas (UAs) of 50,000 or more people, urban clusters (UCs) between 2,500 and 50,000 people, and rural areas (RAs) of less than 2,500 people. It should be noted that USCB uses a different areal unit for urban-rural classification and is not based on the county subdivision (i.e., town). Thus, the category in this analysis does not suggest the actual urban-rural status.

*p-value < 0.05 **p-value < 0.005 ***p-value < 0.001

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
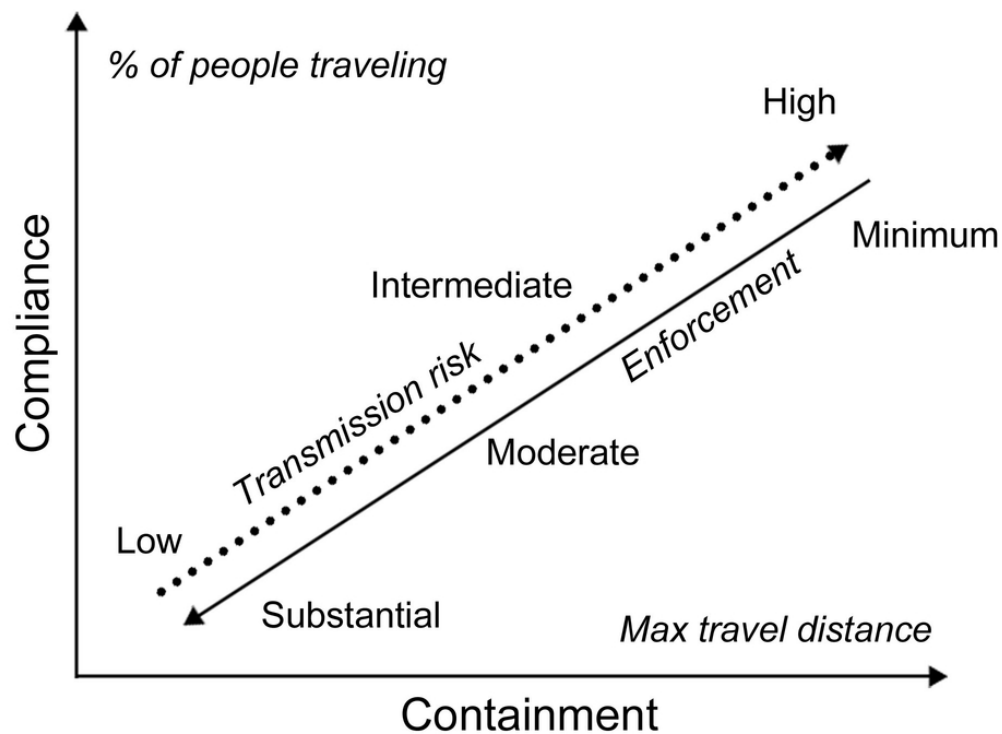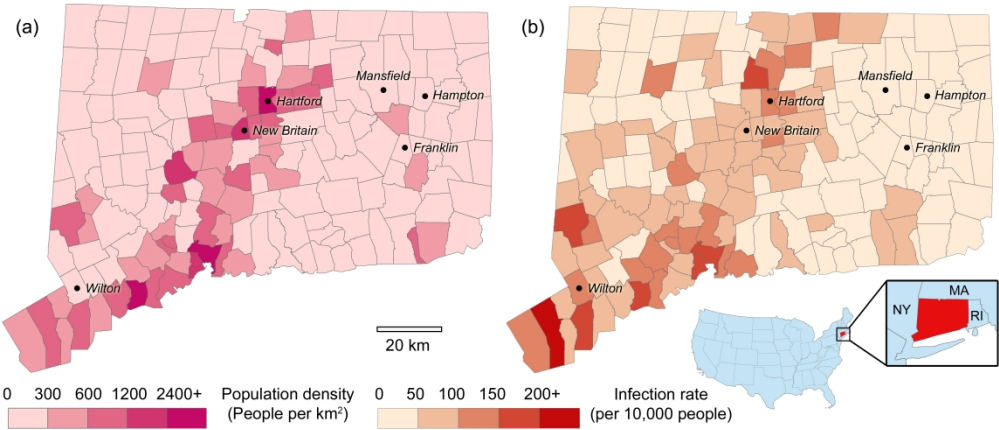45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Figure 1. A conceptual model of the effects of social distancing on travel activities. The solid line represents the strength of policy enforcement; the dashed line represents the level of transmission risk.

77x59mm (300 x 300 DPI)

Figure 2. Connecticut towns with (a) population density and (b) COVID-19 infection rate as of May 11. Towns further discussed in the article are labeled.

390x166mm (300 x 300 DPI)
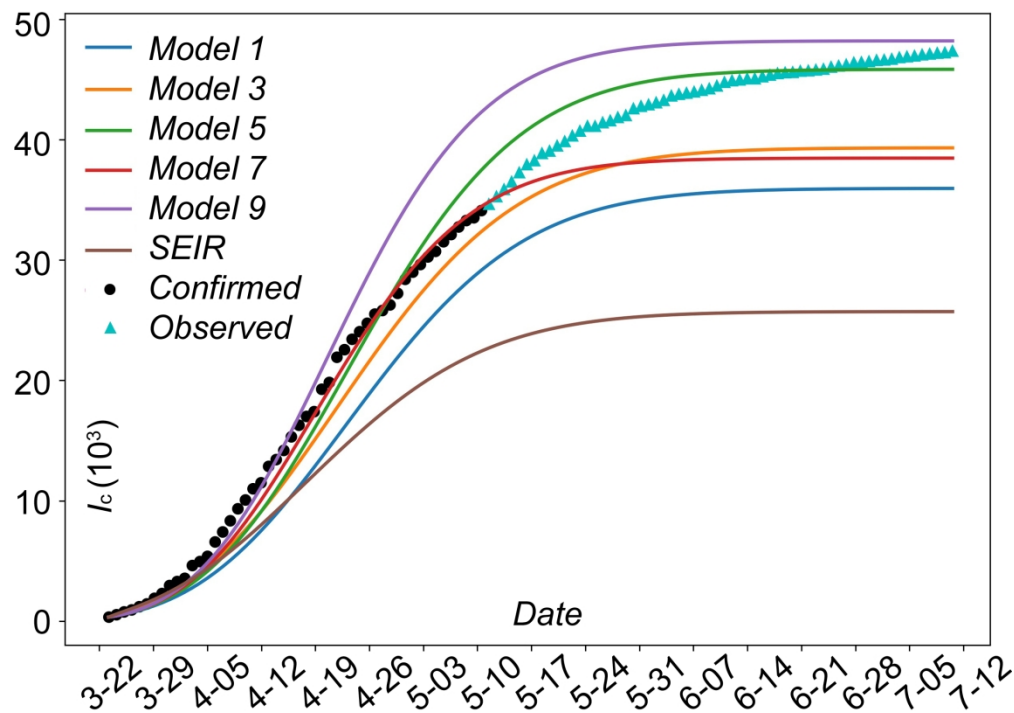
Figure 3. Simulated cumulative cases of infection (Ic) for the state by the SEIR model and the MSEIR model (Models 1, 3, 5, 7, and 9) based on confirmed cases until May 11.
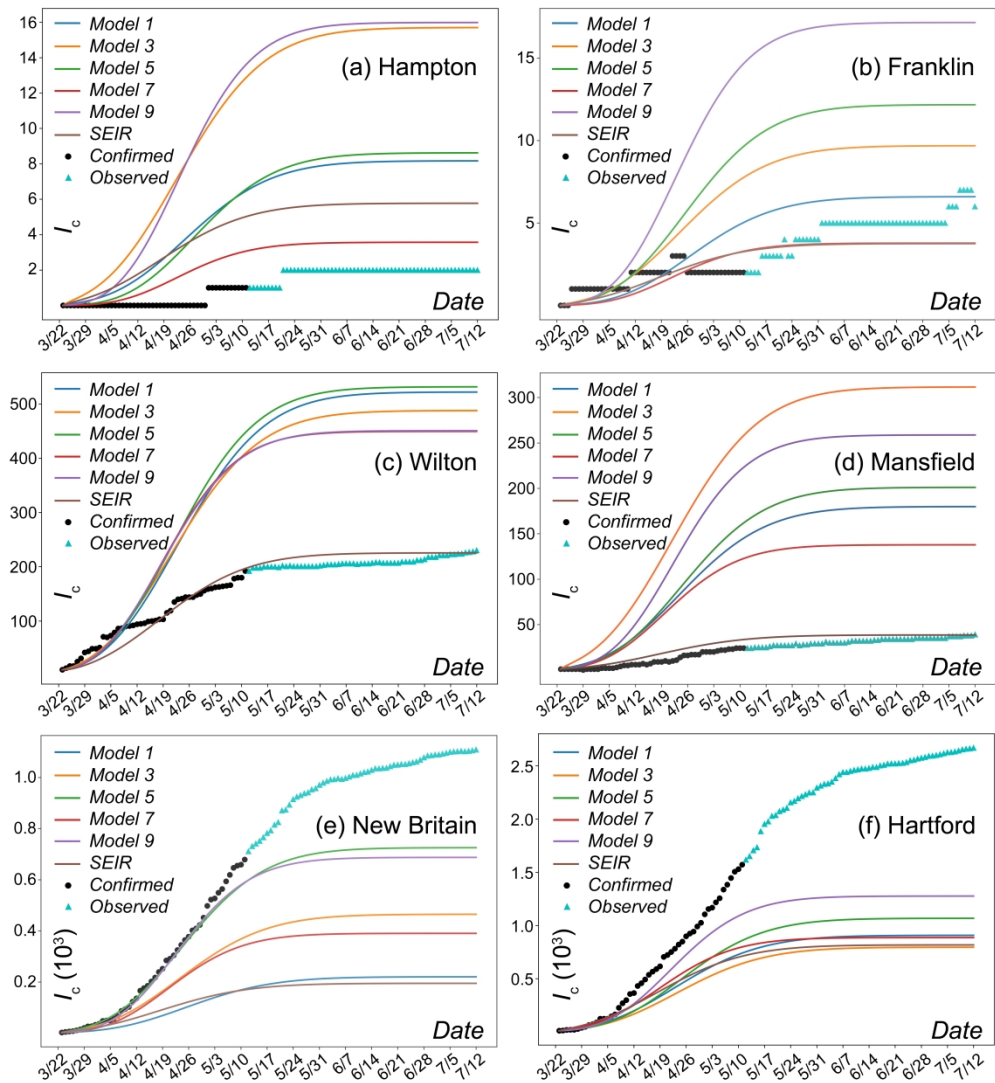
212x149mm (300 x 300 DPI)

Figure 4. Simulated cumulative cases of infection (Ic) by the SEIR model and the MSEIR models for (a) Hampton, (b) Franklin, (c) Wilton, (d) Mansfield, (e) New Britain, and (f) Hartford. Simulations were based on confirmed cases until May 11.
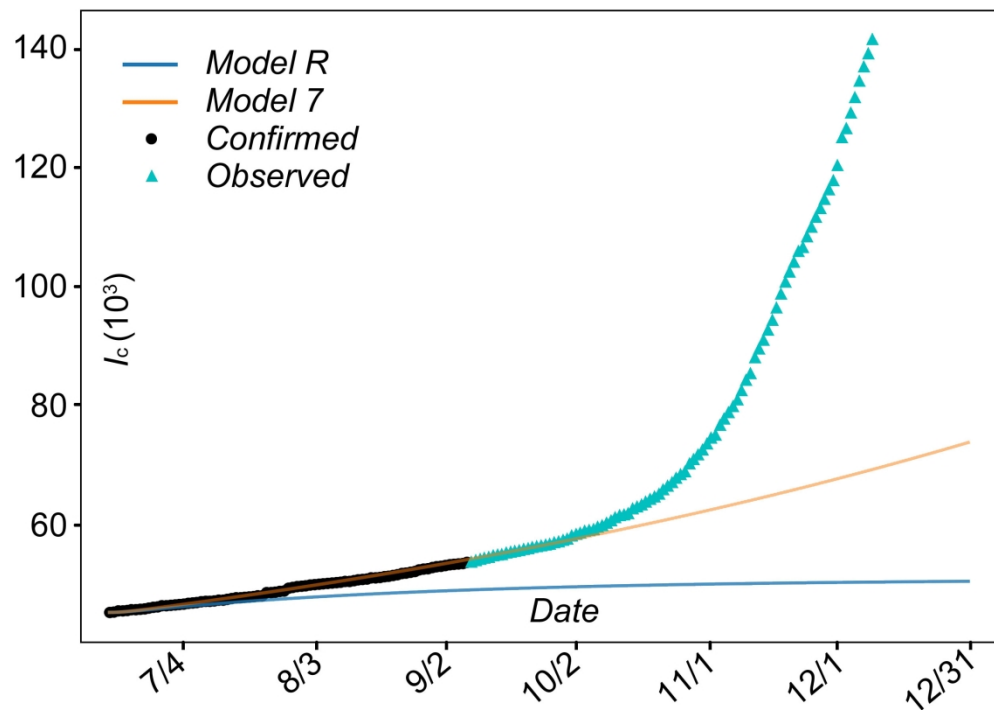
431x465mm (300 x 300 DPI)

Figure 5. Simulated cumulative cases of infection (Ic) for the state by the MSEIR model (Models R and 7) based on confirmed cases until September 7.
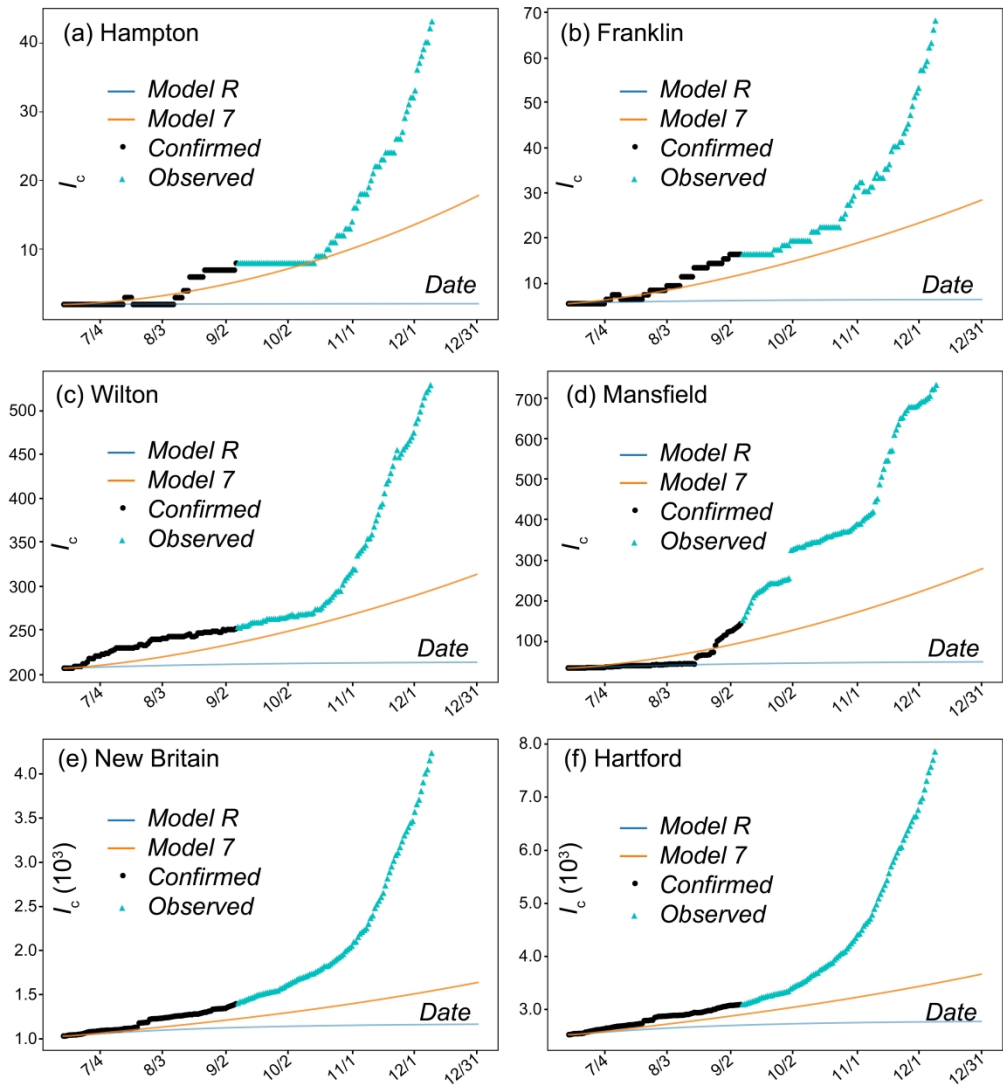
212x149mm (300 x 300 DPI)

Figure 6. Simulated cumulative cases of infection (Ic) by the SEIR model and the MSEIR models for (a) Hampton, (b) Franklin, (c) Wilton, (d) Mansfield, (e) New Britain, and (f) Hartford. Simulations were based on confirmed cases until September 7.

431x465mm (300 x 300 DPI)

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
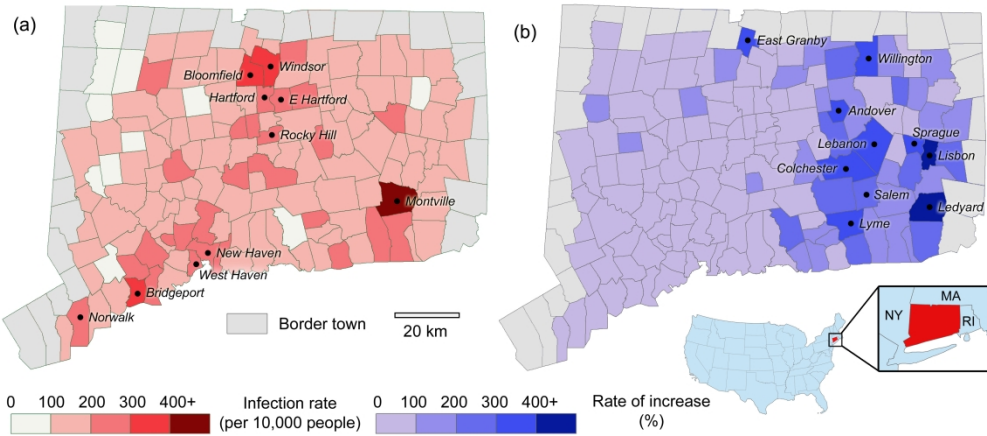51
52
53
54
55
56
57
58
59
60

Figure 7. (1) Simulated infection rate and (2) simulated rate of increase by Model 7 as of December 31. Towns with the top-10 highest rates are labeled.

391x173mm (300 x 300 DPI)