

A Spatio-Temporal Investigation of the Growth Rate of COVID-19 Incidents in Ohio, USA, Early in the Pandemic

ALESSANDRO M. SELVITELLA¹, LIAM CAROLAN, JUSTIN SMETHERS, CHRISTOPHER HERNANDEZ, Department of Mathematical Sciences, Purdue University Fort Wayne, Fort Wayne, IN, USA; and KATHLEEN L. FOSTER, Department of Biology, Ball State University, Muncie, IN, USA.

ABSTRACT. Understanding the initial growth rate of an epidemic is important for epidemiologists and policy makers as it can impact their mitigation strategies such as school closures, quarantines, or social distancing. Because the transmission rate depends on the contact rate of the susceptible population with infected individuals, similar growth rates might be experienced in nearby geographical areas. This research determined the growth rate of cases and deaths associated with COVID-19 in the early period of the 2020 pandemic in Ohio, United States. The evolution of cases and deaths was modeled through a *Besag-York-Mollié* model with linear- and power-type deterministic time dependence. The analysis showed that the growth rate of the time component of the model was subexponential in both cases and deaths once the time-lag across counties of the appearance of the first COVID-19 case was considered. Moreover, deaths in the northeast counties in Ohio were strongly related to the deaths in nearby counties.

Publication Date: July 2021

<https://doi.org/10.18061/ojs.v12i12.8059>

OHIO J SCI 121(2):33-47

INTRODUCTION

In December 2019, a new respiratory virus was first identified in a human patient in China (CDC 2020; WHO 2020). That virus was subsequently classified as *SARS-CoV-2* or *severe acute respiratory syndrome coronavirus 2* and is responsible for the disease now commonly referred to as *COVID-19* that became a global pandemic in 2020 (WHO 2020).

A major concern with any virus is how fast it spreads, and in particular the growth rate of cases and deaths. Growth rates are especially important in the initial stages of a pandemic, when they are usually exponential (Diekmann et al. 2013; Chowell et al. 2016; Brauer et al. 2019; Hilton and Keeling 2020). Because the transmission rate depends on the contact rate of the susceptible population with infected individuals, similar growth rates might be experienced in nearby geographical areas. Geographical regions might have region-specific factors which affect the growth rate, such as population density, mobility of the population, access to healthcare, and a number of other factors (Selvitella and Foster 2020; Foster and Selvitella 2021).

Having an understanding of the spread of COVID-19 early in the pandemic is important for multiple reasons. A fast spread can put a strain on the healthcare system, especially hospitals which may be forced to work either at or beyond their capacity (as they were in the first months of the pandemic (Miller et al. 2020)). Shortages of medical and protective equipment, combined with government mandated shutdowns of the economy, can occur in such times of high healthcare burden (Miller et al. 2020; Selvitella and Foster 2020; Foster and Selvitella 2021). Having information on the spatial and temporal characteristics of the transmission of COVID-19 during the first stages of the pandemic was of assistance to policy makers (who strive for a more effective allocation of limited resources) and helped to predict the dynamics of early stages of subsequent waves or novel mutations.

The interest of this study lies in understanding the spatio-temporal diffusion of COVID-19, specifically understanding the spread of COVID-19 across counties in Ohio. A *Besag-York-Mollié* (BYM) model (Besag et al. 1991; Blangiardo and Cameletti 2015) was developed from the cases and deaths in Ohio for the first 3 weeks following the initial

¹ Address correspondence to Alessandro Maria Selvitella, Department of Mathematical Sciences, Purdue University Fort Wayne, 267 Kettler Hall, 2101 E. Coliseum Blvd., Fort Wayne, IN 46805, USA. Email: aselvite@pfw.edu



COVID-19 case in each county. The first county in which the virus was detected was Cuyahoga on 9 March 2020; the last day considered in the study was 5 May 2020 (in Harrison County).

This model provides insight to the following questions: First, to what extent does spatial correlation influence the growth of COVID-19 cases and deaths in Ohio counties? Second, was the initial growth of cases in Ohio subexponential, exponential, or superexponential (Tolle 2003)? This second question is important for both epidemiologists and policymakers. Indeed, an important factor to determine when a pandemic begins is the initial growth rate, which in turn gives information to the most important epidemiological parameter, the *basic reproduction number*—the expected number of secondary cases caused by a single infected individual in a fully susceptible population (Diekmann et al. 2013; Brauer et al. 2019). In the first weeks of pandemics, the main component of the growth curve of fatalities is generally of exponential type, but in some cases it has been shown to have a subexponential (Chowell et al. 2016) or superexponential (Tolle 2003) behavior.

The paper is completed with: Appendix A providing relevant geographical information about Ohio, Appendix B illustrating the R-code developed, and Appendix C containing the estimates of the relevant parameters of the models developed (in particular those used for the Results and Discussion Section).

METHODS

Dataset and Software

The dataset analyzed (CDC 2020) contained the number of cases and deaths associated with COVID-19 in the first 3 weeks of the pandemic, beginning with the first day COVID-19 appeared in each of the 88 counties in Ohio. Note that the first case did not appear on the same day in every county. For data analytic purposes, 3 supplemental spreadsheets titled "Supplemental Table 1," "Supplemental Table 2," and "Supplemental Table 3" are available at: <http://hdl.handle.net/1811/92834>.

The statistical analysis utilized the R-INLA statistical package (Rue et al. 2009; R-INLA 2020) available for R statistical software (R Core Team 2020). R-INLA is a package that performs approximate Bayesian inference for latent Gaussian models such as those considered in this manuscript.

Spatio-Temporal Model

Suppose y_{it} represents the number of cases or deaths at a discrete time t with $t=1, \dots, T$ with T the time extension of this analysis. The index $i=1, \dots, n$ identifies the county in Ohio with n the total number of counties in Ohio. It is assumed that $y_{it} \sim \text{Poiss}(\lambda_{it})$, namely the case where the cases and deaths are distributed with a Poisson random variable with parameter λ_{it} is considered. Note that this assumption implies that the cases and deaths were equidispersed (McCullagh and Nelder 1989). The parameter λ_{it} is assumed to be the product $\lambda_{it} = E_i \rho_{it}$. While ρ_{it} represents the rate of appearance of new cases and deaths at time t in area (county) i , the expected number of cases or deaths in area i is defined as follows. Suppose the population under consideration is partitioned into J classes and r_j is the standardized reference rate for class $j=1, \dots, J$ and P_{ij} is the population count of class j in county i . Then the expected number of cases or deaths in area i is given by:

$$E_i = \sum_{j=1}^J P_{ij} r_j.$$

A link function $\eta_{it} = \log(\rho_{it})$ is used to relate the cases or deaths to their spatio-temporal covariates in the following way:

$$\log(\rho_{it}) = \eta_{it} = b_0 + u_i + v_i + \Phi_{it}.$$

The intercept b_0 represents the average outcome rate in the entire study region (i.e., Ohio in this case); the covariate v_i represents the area-specific effect and it is modeled as exchangeable, while to u_i is given an autoregressive structure (Besag 1974; Blangiardo and Cameletti 2015). More specifically, suppose that $N(i)$ represents the set of neighborhoods of area i , for $i=1, \dots, n$ and that u_{-i} represents the area-specific effects excluding county i . Then, the conditional distribution $u_i | u_{-i}$ is given by:

$$u_i | u_{-i} \sim N(\psi_i, \theta_i^2),$$

with

$$\psi_i = \mu_i + \sum_{k=1}^n r_{ik} (u_k - \mu_k) \text{ and } \theta_i^2 = s_i^2$$

where μ_i represents the mean for area i , $s_i^2 = \sigma_u^2 / N_i$ is the variance of area i , and $N_i = |N(i)|$ is the number of neighborhoods of area i . In this formulation, σ_u^2 represents the amount of variation between the spatially-structured random effects, and r_{ik} is the

indicator of spatial proximity between areas i and k which is defined as $r_{ik} = \phi W_{ik}$ with $W_{ik} = a_{ik}/N_i$ and $a_{ik} = 1$ if i and k are neighbors and 0 otherwise. The factor $\phi > 0$ is introduced to control the properness of the probability distribution and needs to ensure the positive definiteness of the variance-covariance matrix $(I - \phi W) S^2$, with I the $n \times n$ identity matrix, $W = \{W_{ik}\}_{i,k=1}^n$, and $S^2 = \text{diag}\{s_1^2, \dots, s_n^2\}$ (Cressie 1993). With these assumptions, the proper *conditional autoregressive* (CAR) specification u is a multivariate normal random variable, defined as follows:

$$u \sim MVN(\mu, (I - \phi W)S^2),$$

with $\mu = \{\mu_1, \dots, \mu_n\}$ the mean vector. This structure implies that the correlation between areas i and j can be computed using the following formula (Blangiardo and Cameletti 2015):

$$\frac{\sqrt{N_i}}{N_k} \frac{(I - \phi W)_{ik}^{-1}}{\sqrt{(I - \phi W)_{ii}^{-1} (I - \phi W)_{kk}^{-1}}}$$

It is further assumed that $\phi = 1$ and

$$\sum_{i=1}^n u_i = 0,$$

a specification called *intrinsic conditional autoregressive* which together with the exchangeable random effect gives rise to the so-called Besag-York-Mollié model (BYM) (Besag et al. 1991; Blangiardo and Cameletti 2015). By assuming that $\mu_i = 0$ for every $i = 1, \dots, n$, then

$$u_i | u_{-i} \sim N\left(\frac{1}{N_i} \sum_{k=1}^n a_{ik} u_k, s_i^2\right).$$

(Refer to Besag et al. 1991; Best et al. 2005; Lee 2011; and Lawson 2018 for further details.)

Concerning the temporal component Φ_{it} , a parametric trend is considered (as presented in Bernardinelli et al. 1995) which assumes

$$\Phi_{it} = (\beta + \delta_i)t^p$$

for $t = 1, \dots, T$ and $i = 1, \dots, n$, but with p to be determined and not necessarily $p = 1$. Here, β represents the main trend, while δ_i represents the differential trends and the interaction between space and time components with the identifiability assumption

$$\sum_{i=1}^n \delta_i = 0.$$

Deviance Information Criterion

The *deviance information criterion* (DIC) is a measure of model fitting typically used for Bayesian models (Spiegelhalter et al. 2002) which includes a trade-off between goodness of fit and model complexity.

Given a random variable Y , whose distribution depends on a set of parameters θ , the DIC is given by the following formula:

$$DIC = \underline{D} + p_D.$$

In this formula, p_D represents the *effective number of parameters*, which is computed as

$$p_D = \underline{D} - D(\underline{\theta})$$

with

$$\underline{D} := E_{\theta|y}[D(\theta)] \text{ and } D(\underline{\theta}) := D(E_{\theta|y}[\theta]).$$

Here, $D(\theta) := -2\log(p(y|\theta))$ represents the *deviance* and \underline{D} the posterior expectation of the deviance. The DIC can be equivalently rewritten as

$$DIC = D(\underline{\theta}) + 2p_D,$$

in which case it resembles the common expression of the *Akaike information criterion* (AIC), of which the DIC is a generalization (James et al. 2013).

Statistical Analysis

The statistical analysis is divided into 2 parts. In both cases, homogeneity in the population is assumed, namely that $E_i = 1$ for every $i = 1, \dots, n$.

In the first part of the study, the cases and deaths were modeled using the linear ($p = 1$) BYM and information about the spatial correlation and temporal differential effect emerging from the assumptions was extracted. It is worth noticing that, in these models, the time index is translated of a county-dependent factor t_i , $i = 1, \dots, n$ so that the analysis for each county starts the day in which the first case appeared in that county. This procedure was performed before modeling.

These results are visualized by plotting both the *posterior mean of the spatial main effect* $\zeta_i = \exp(u_i + v_i)$ and the *differential effect* δ_i , for $i = 1, \dots, n$ of the cases and deaths models with linear growth rate $p = 1$. The thresholds which determine the classes (colors) in the plots are based on the $0, 0.25, 0.5, 0.75, 1$ -quantiles of the empirical distributions of the corresponding parameter for all Ohio counties.

In the second part of the study, the best models were determined according to the DIC among the family of BYM with a generic nonlinear term $p \in [0,2]$, with the search performed over steps of length 0.1. The search for an optimal model was performed independently and for both outcome variables (number of cases and number of deaths).

RESULTS AND DISCUSSION

Spatial Main Effect and Temporal Differential Effect in the Linear Model

Fig. 1 and Fig. 2 represent the distribution of spatial and temporal effects, respectively, on the COVID-19 cases by county in Ohio for the linear model $p = 1$.

The spatial effect is given by the posterior mean of the spatial main effect $\zeta_i = \exp(u_i + v_i)$ for $i = 1, \dots, n$. For the full output, see Appendix C. The counties

with the darkest color (red) are the counties where the spatial effect ζ_i is stronger (Fig. 1). Especially in northeast Ohio and southwest Ohio, the model supports the presence of spatial correlation between counties.

The temporal effect is given by the differential effect δ_i for $i = 1, \dots, n$ and represents the extra growth rate of county i with respect to the Ohio growth rate. For the full output, see Appendix C. The counties with the darkest color (red) are the counties where the differential temporal effect δ_i is stronger (Fig. 2). Again, in northeast and southwest Ohio, the model mildly supports the presence of the temporal differential effect.

Fig. 3 and Fig. 4 represent the distribution of spatial and temporal effects, respectively, on the COVID-19 related deaths by county in Ohio for the linear model $p = 1$.

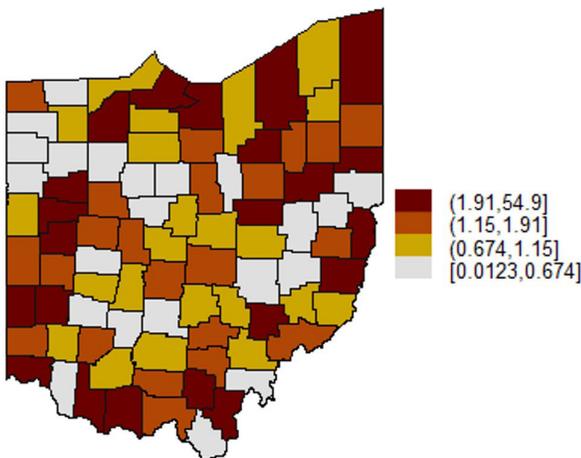


FIGURE 1. Posterior mean of the spatial main effect $\zeta_i = \exp(u_i + v_i)$ for $i = 1, \dots, n$ of the cases model with linear growth rate power $p = 1$

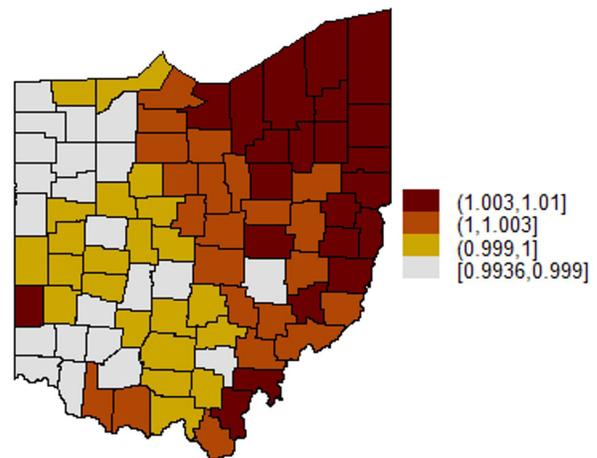


FIGURE 3. Posterior mean of the spatial main effect $\zeta_i = \exp(u_i + v_i)$ for $i = 1, \dots, n$ of the deaths model with linear growth rate power $p = 1$

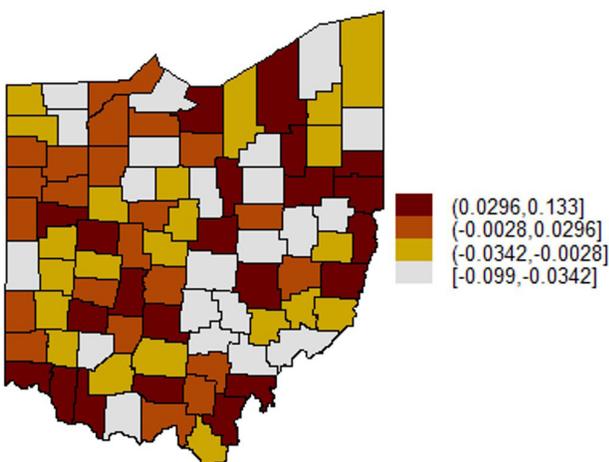


FIGURE 2. Temporal differential effects δ_i for $i = 1, \dots, n$ of the cases model with linear growth rate power $p = 1$

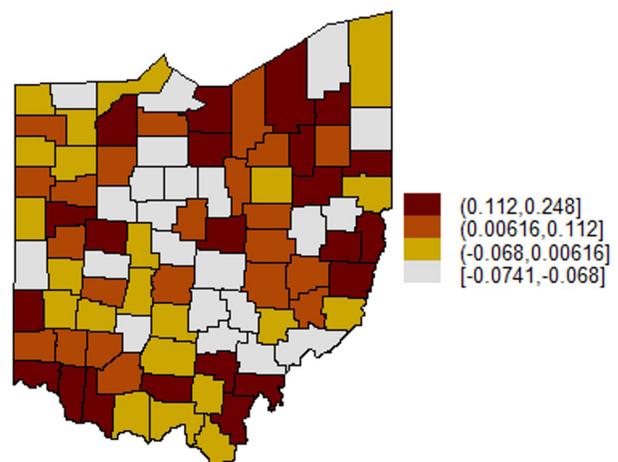


FIGURE 4. Temporal differential effects δ_i for $i = 1, \dots, n$ of the deaths model with linear growth rate power $p = 1$

In Fig. 3, the spatial effect is given by the posterior mean of the spatial main effect $\zeta_i = \exp(u_i + v_i)$ for $i = 1, \dots, n$. For the full output, see Appendix C. Also, the counties with the darkest color (red) are the counties where the spatial effect ζ_i is stronger. The spatial effect in this case is striking and shows a strong spatial correlation of the diffusion of COVID-19 in the first 3 weeks of the pandemic. Northeast Ohio, in particular, is the area where the posterior mean of the spatial main effect is the strongest, with gradual decrease in the radial direction from northeast to southwest.

In Fig. 4, the temporal effect is given by the differential effect δ_i for $i = 1, \dots, n$ and represents the extra growth rate of county i with respect to the Ohio growth rate. For the full output, see Appendix C. The counties with the darkest color (red) are the counties where the differential temporal effect δ_i is stronger. Similar to the model for the cases, in northeast Ohio and southwest Ohio, the model mildly supports the presence of temporal differential effect.

It is important to remember that the 21 days considered for each county are not temporally aligned, as the 21 days are considered from the day the first case of COVID-19 was reported in that county. This means that the temporal component due to the lag between the first case in one county and the first in another county, which is substantial, has already been taken into account in both models (number of both cases and deaths). This is in agreement with the milder effect of the temporal component in the models. Nevertheless, this analysis shows that even quotienting out the natural differential time-lag mentioned above, there is still some residual temporal effect distinguishable across counties.

Optimal Growth

This statistical analysis shows that the best model according to the DIC has sublinear growth both in the models where the outcome variables are the number of cases and those where the outcome variables are the number of deaths.

In particular, the model for the cases has optimal time growth when the exponent of the time nonlinearity is $p = 0.3$ (Table 1).

The model for the deaths has optimal time growth when the exponent of the time nonlinearity is $p = 0.7$ (Table 2).

Qualitatively, this indicates that the residual time component—once the differential time-lag of the appearance of the first case across counties had been accounted—had a sublinear growth. Understanding the exact growth rate of the number of cases and deaths has crucial implications for public health and can guide governmental policy decisions. Note that the analysis performed in this study includes a single family of models, so it cannot be used alone to guide a governmental decision on such an important issue.

Conclusions

This research studied the growth rate of the incidents (both cases and deaths) associated with COVID-19, in Ohio, during the 21 days after the first case appeared in each county. The evolution of the cases and deaths was modeled through a Besag-York-Mollié model with linear- and power-type time dependence. This model showed that the spatial effect on the deaths was substantial throughout Ohio, while the residual differential effect of the time component was mild, but present, in every county. In particular, the northeast counties in Ohio exhibited the strongest spatial effect, especially in the model of deaths related to COVID-19. Furthermore, this analysis showed that the growth rate of cases and deaths in Ohio were subexponential in the early stages of the pandemic—once the natural lag between the appearance of the first case per county was taken into consideration.

As more data become available, future analysis will be dedicated to comparing the initial growth rate to the growth rate of cases and deaths at the beginning of subsequent waves of the pandemic and to determine the possibility of seasonal components. The initial growth rate of an epidemic relates to the basic reproduction number and so the former is of great importance to policy makers as it can influence governmental decisions and prompt mitigation strategies.

ACKNOWLEDGEMENTS

The authors would like to thank the editor and the anonymous reviewers for the careful reading of our manuscript.

Table 1
DIC values for increasing powers of p for the model of the cases

p	DIC
0	15183.75
0.1	11266.22
0.2	11186.54
0.3	11184.46
0.4	11211.81
0.5	11258.37
0.6	11319.57
0.7	11392.70
0.8	11475.93
0.9	11568.46
1.0	11670.49
1.1	11782.90
1.2	11907.54
1.3	12046.86
1.4	12204.18
1.5	12383.50
1.6	12589.55
1.7	12827.68
1.8	13103.30
1.9	13421.04
2.0	13783.03

Table 2
DIC values for increasing powers of p for the model of the deaths

p	DIC
0	6.681565e+48
0.1	1.748291e+03
0.2	1.745994e+03
0.3	1.745208e+03
0.4	1.744214e+03
0.5	1.743577e+03
0.6	1.743111e+03
0.7	1.743089e+03
0.8	1.743577e+03
0.9	1.744467e+03
1.0	1.746373e+03
1.1	1.749216e+03
1.2	1.755190e+03
1.3	1.775394e+03
1.4	1.095150e+84
1.5	1.897523e+69
1.6	5.012129e+63
1.7	4.404118e+61
1.8	1.392719e+61
1.9	7.181979e+60
2.0	2.166742e+60

LITERATURE CITED

- Bernardinelli L, Clayton D, Pascutto C, Montomoli C, Ghislandi M, Songini M. 1995. Bayesian analysis of space-time variation in disease risk. *Stat Med.* 14(21-22):2433-2443. <https://doi.org/10.1002/sim.4780142112>
- Besag J. 1974. Spatial interaction and the statistical analysis of lattice systems. *J Roy Stat Soc B Met.* 36(2):192-236 <http://www.jstor.org/stable/2984812>
- Besag J, York J, Mollie A. 1991. Bayesian image restoration, with two applications in spatial statistics. *Ann I Stat Math.* 43:1-20. <https://doi.org/10.1007/BF00116466>
- Best N, Richardson S, Thomson A. 2005. A comparison of Bayesian spatial models for disease mapping. *Stat Methods Med Res.* 14(1):35-59. <https://doi.org/10.1191/0962280205sm388oa>
- Blangiardo M, Cameletti M. 2015. Spatial and spatio-temporal Bayesian models with R-INLA. Chichester, West Sussex (UK): John Wiley & Sons Ltd. Chapter 5, Bayesian regression and hierarchical models; p.127-172. ISBN 13: 9781118326558. <https://doi.org/10.1002/9781118950203.ch5>
- Brauer F, Castillo-Chavez C, Feng Z. 2019. Mathematical models in epidemiology. New York (NY): Springer. 619 p. ISBN 13: 9781493998265. <https://doi.org/10.1007/978-1-4939-9828-9>
- [CDC] Center for Disease Control and Prevention. 2020. Atlanta (GA): United States Department of Health & Human Services; [accessed 2020 December 21]. <https://www.cdc.gov>
- Chowell G, Viboud C, Simonsen L, Moghadas SM. 2016. Characterizing the reproduction number of epidemics with early subexponential growth dynamics. *J Roy Soc Interface.* 13(123):20160659. <https://doi.org/10.1098/rsif.2016.0659>
- Cressie NAC. 1993. Statistics for spatial data: revised edition. New York (NY): John Wiley & Sons Inc. 928 p. ISBN 13: 9781119115151. <https://doi.org/10.1002/9781119115151>
- Diekmann O, Heesterbeek H, Britton T. 2013. Mathematical tools for understanding infectious disease dynamics. Princeton (NJ): Princeton University Press. 502 p. ISBN 13: 9780691155395. <https://doi.org/10.1515/9781400845620>
- Foster KL, Selvitella AM. 2021. On the relationship between COVID-19 reported fatalities early in the pandemic and national socio-economic status predating the pandemic. *AIMS Public Health.* 8(3):439-455. <https://doi.org/10.3934/publichealth.2021034>
- Hilton J, Keeling MJ. 2020. Estimation of country-level basic reproductive ratios for novel coronavirus (SARS-CoV-2/COVID-19) using synthetic contact matrices. *Plos Comput Biol.* 16(7):e1008031. <https://doi.org/10.1371/journal.pcbi.1008031>
- James G, Witten D, Hastie T, Tibshirani R. 2013. An introduction to statistical learning. New York (NY): Springer. 426 p. ISBN 13: 9781461471370. <https://doi.org/10.1007/978-1-4614-7138-7>
- Lawson AB. 2018. Bayesian disease mapping: hierarchical modeling in spatial epidemiology. 3rd ed. Boca Raton (FL): Chapman and Hall/CRC. 464 p. ISBN 13: 9781351271769. <https://doi.org/10.1201/9781351271769>
- Lee D. 2011. A comparison of conditional autoregressive models used in Bayesian disease mapping. *Spatial Spatio Temporal Epidemiol.* 2(2):79-89. <https://doi.org/10.1016/j.sste.2011.03.001>
- McCullagh P, Nelder JA. 1989. Generalized linear models. 2nd ed. Boca Raton (FL): Chapman & Hall/CRC. 532 p. ISBN 13: 9780412317606.
- Miller IF, Becker AD, Grenfell BT, Metcalf CJE. 2020. Disease and healthcare burden of COVID-19 in the United States. *Nat Med.* 26:1212-1217. <https://doi.org/10.1038/s41591-020-0952-y>
- R Core Team. 2020. R: a language and environment for statistical computing. Vienna (Austria): R Foundation for Statistical Computing. <https://www.r-project.org/>
- [R-INLA]. R-INLA project software package homepage. 2020. [accessed 2020]. <https://www.r-inla.org/home>
- Rue H, Martino S, Chopin N. 2009. Approximate Bayesian inference for latent Gaussian models by using integrated nested Laplace approximations. *J Roy Stat Soc B.* 71(2):319-392. <https://doi.org/10.1111/j.1467-9868.2008.00700.x>
- Selvitella AM, Foster KL. 2020. Societal and economic factors associated with COVID-19 indicate that developing countries could suffer the most. *Technium Soc Sci J.* 10(1):637-644. <https://doi.org/10.47577/tssj.v10i1.1357>
- Spiegelhalter DJ, Best NG, Carlin BP, van der Linde A. 2002. Bayesian measures of model complexity and fit. *J Roy Stat Soc B.* 64(4):583-639. <https://doi.org/10.1111/1467-9868.00353>
- Tolle J. 2003. 87.65 Can growth be faster than exponential, and just how slow is the logarithm? *Math Gaz.* 87(510):522-525. <https://www.jstor.org/stable/3621297>
- [WHO] World Health Organization. 2020. Rolling updates on coronavirus disease (COVID-19) [event timeline]. [date updated 2020 July 31]. <https://www.who.int/emergencies/diseases/novel-coronavirus-2019/events-as-they-happen>

APPENDICES

Appendix A: Ohio Geography

Editor's Note: The authors assert, and the reviewers acknowledge, that the discrepancy between county names in the map and the county names read internally to the R-code below did not affect the overall analysis. Specifically, the names in the map are correct, but the names of the columns in the R-code below include the county name "Unknown" and omit Wyandot County. This inconsistency is derived from the original Excel file downloaded from the CDC that is posted here (<http://hdl.handle.net/1811/92834>) and whose column names were not modified during the analysis and kept as read by the command "Import Dataset" in the R-software.

For reference, a map of Ohio with the names of all 88 counties is provided. This can be used to identify the specific county from the plot reported in previous sections and specific parameters as reported in Appendix C below.



Names of all 88 counties in Ohio

Below are the names of the counties in alphabetical order.

```
unique(data$`county name`)
[1] "Adams" "Allen" "Ashland" "Ashtabula" "Athens" "Auglaize"
[7] "Belmont" "Brown" "Butler" "Carroll" "Champaign" "Clark"
[13] "Clermont" "Clinton" "Columbiana" "Coshocton" "Crawford" "Cuyahoga"
[19] "Darke" "Defiance" "Delaware" "Erie" "Fairfield" "Fayette"
[25] "Franklin" "Fulton" "Gallia" "Geauga" "Greene" "Guernsey"
[31] "Hamilton" "Hancock" "Hardin" "Harrison" "Henry" "Highland"
[37] "Hocking" "Holmes" "Huron" "Jackson" "Jefferson" "Knox"
[43] "Lake" "Lawrence" "Licking" "Logan" "Lorain" "Lucas"
[49] "Madison" "Mahoning" "Marion" "Medina" "Meigs" "Mercer"
[55] "Miami" "Monroe" "Montgomery" "Morgan" "Morrow" "Muskingum"
[61] "Noble" "Ottawa" "Paulding" "Perry" "Pickaway" "Pike"
[67] "Portage" "Preble" "Putnam" "Richland" "Ross" "Sandusky"
[73] "Scioto" "Seneca" "Shelby" "Stark" "Summit" "Trumbull"
[79] "Tuscarawas" "Union" "Unknown" "Van Wert" "Vinton" "Warren"
[85] "Washington" "Wayne" "Williams" "Wood"
```

Each index corresponds to the indexes in the codes and estimates in Appendix B and Appendix C below.

Appendix B: R-Code

This appendix is devoted to the code used for the analysis, for reproducibility purposes.

This part of the code loads the packages and the dataset.

```
#Set Directory
remove(list=ls())
setwd("XX")
my.dir <- paste(getwd(), "/", sep="")

#Require Packages
require(INLA)
inla.setOption(scale.model.default=FALSE)
require(splancs)
require(sp)
require(fields)
require(lattice)
require(rgdal)
require(abind)

#Upload the Dataset
library(readxl)
CountyCovid_OhioOnly_ <- read_excel("XX/CountyCovid(OhioOnly)_days.xlsx")
data<-CountyCovid_OhioOnly_[1:15687,]
```

Here only the data from the first 21 days since the appearance of the first case in each county is selected.

```
index<-vector()
n=88
for (i in 1:n) {
  print(i)
  #data[which(data$county==i)[1:21],]
  index<-c(index,which(data$county==i)[1:21])
}

data21<-data[index,]

Ohio <- readOGR(paste(my.dir, "XX/tl_2010_39_county00.shp", sep=""))
county1 <- data21$county
Ohio.adj <- paste(my.dir, "XX/Ohio.graph", sep="")
```

This code constructs the Linear BYM model for the number of cases and produces Fig. 1 and Fig. 2.

```
formula.par <- cases ~ 1 + f(data21$county,model="bym",graph=Ohio.adj, constr=TRUE)
+ f(county1,time,model="iid", constr=TRUE)+ time

model.par <- inla(formula.par,family="poisson",data=data21,
  control.predictor=list(compute=TRUE),
  control.compute=list(dic=TRUE,cpo=TRUE))
round(model.par$summary.fixed[,1:5],3)

""

      mean  sd 0.025quant 0.5quant 0.975quant
(Intercept) -1.030 0.144   -1.316  -1.029   -0.750
I(time^1)   0.124 0.002    0.121  0.124    0.128
```

```

'''

model.par$dic$dic
#11670.49

# Figure 1
m <- model.par$marginals.random[[1]][1:88]
zeta.ST1 <- unlist(lapply(m,function(x)inla.emarginal(exp,x)))

SMR.cutoff<-quantile(zeta.ST1, probs = c(0,0.25,0.5,0.75,1))
SMR.cutoff
xi.factor <- cut(zeta.ST1,breaks=SMR.cutoff,include.lowest=TRUE)
m <- model.par$summary.random[[2]][1:88,2]
#int.cut <- c(-0.008,-0.001,0,0.001,0.006)
int.cut<-quantile(m, probs = c(0,0.25,0.5,0.75,1))
int.factor <- cut(m,breaks=int.cut,include.lowest=TRUE)

data.ohio <- attr(Ohio, "data")
attr(Ohio, "data") <- data.frame(data.ohio, xi=xi.factor)
trellis.par.set(axis.line=list(col=NA))
spplot(obj=Ohio, zcol= "xi", col.regions=gray(3.5:0.5/4),main="")

#Figure 2
data.ohio2 <- attr(Ohio, "data")
attr(Ohio, "data") <- data.frame(data.ohio2, int=int.factor)
spplot(obj=Ohio, zcol= "int",col.regions=gray(3.5:0.5/4),main="")

This code constructs the Linear BYM model for the number of deaths and produces Fig. 3 and Fig. 4.
formula2.par <- deaths ~ 1 + f(data21$county,model="bym",graph=Ohio.adj, constr=TRUE)
+ f(county1,time,model="iid", constr=TRUE)+ time

model2.par <- inla(formula2.par,family="poisson",data=data21,
  control.predictor=list(compute=TRUE),
  control.compute=list(dic=TRUE,cpo=TRUE))
round(model2.par$summary.fixed[,1:5],3)

'''

      mean  sd 0.025quant 0.5quant 0.975quant
(Intercept) -4.550 0.210   -4.968  -4.547  -4.145
time         0.054 0.013    0.028   0.055   0.078
'''

model2.par$dic$dic
#1746.373

#Figure 3
m2 <- model2.par$marginals.random[[1]][1:88]
zeta.ST1_2 <- unlist(lapply(m2,function(x)inla.emarginal(exp,x)))

```

```

SMR.cutoff_2<-quantile(zeta.ST1_2, probs = c(0,0.25,0.5,0.75,1))
SMR.cutoff_2
xi.factor_2 <- cut(zeta.ST1_2,breaks=SMR.cutoff_2,include.lowest=TRUE)
xi.factor_2
m2 <- model2.par$summary.random[[2]][1:88,2]

int.cut_2<-quantile(m2, probs = c(0,0.25,0.5,0.75,1))
int.cut_2
int.factor_2 <- cut(m2,breaks=int.cut_2,include.lowest=TRUE)
int.factor_2

data.ohio <- attr(Ohio, "data")
attr(Ohio, "data") <- data.frame(data.ohio, xi=xi.factor_2)
trellis.par.set(axis.line=list(col=NA))
splot(obj=Ohio, zcol= "xi", col.regions=gray(3.5:0.5/4),main="")

# Figure 4
data.ohio2 <- attr(Ohio, "data")
attr(Ohio, "data") <- data.frame(data.ohio2, int=int.factor_2)
splot(obj=Ohio, zcol= "int",col.regions=gray(3.5:0.5/4),main="")

```

This R-code outputs the DIC for the optimal growth rate for the model of the number of cases.

```

DIC<-vector()

v=seq(0,2,0.1)
for (p in v){

formula.par <- cases ~ 1 + f(data21$county,model="bym",graph=Ohio.adj, constr=TRUE)
+ f(county1,time,model="iid", constr=TRUE)+ I(time^p)

model.par <- inla(formula.par,family="poisson",data=data21,
                  control.predictor=list(compute=TRUE),
                  control.compute=list(dic=TRUE,cpo=TRUE))
round(model.par$summary.fixed[,1:5],3)
DIC<-c(DIC, model.par$dic$dic)
print(model.par$dic$dic)
}

#DIC vector
print(DIC)
""
[1] 15183.60 11266.22 11186.54 11184.46 11211.81 11258.37 11319.57 11392.69 11475.93
[10] 11568.46 11670.49 11782.90 11907.54 12046.86 12204.18 12383.50 12589.55 12827.68
[19] 13103.30 13421.04 13783.03
""

#Best DIC value
print(min(DIC))
""
11184.46
""

#Best DIC model
print(which.min(DIC))

```

```

#4 p=0.3

###
#Best Model for Cases has p=0.3
###

p=0.3
formula.par <- cases ~ 1 + f(data21$county,model="bym",graph=Ohio.adj, constr=TRUE)
+ f(county1,time,model="iid", constr=TRUE)+ I(time^p)

model.par <- inla(formula.par,family="poisson",data=data21,
  control.predictor=list(compute=TRUE),
  control.compute=list(dic=TRUE,cpo=TRUE))
round(model.par$summary.fixed[,1:5],3)

""
      mean  sd 0.025quant 0.5quant 0.975quant
(Intercept) -8.762 0.237  -9.232  -8.760  -8.301
I(time^p) 4.105 0.057  3.994  4.105  4.216

```

This R-code outputs the DIC for the optimal growth rate for the model of the number of deaths.

```

DIC_2<-vector()
v=seq(0,2,0.1)
for (p in v){

  formula2.par <- deaths ~ 1 + f(data21$county,model="bym",graph=Ohio.adj, constr=TRUE)
  + f(county1,time,model="iid", constr=TRUE)+ I(time^p)

  model2.par <- inla(formula2.par,family="poisson",data=data21,
    control.predictor=list(compute=TRUE),
    control.compute=list(dic=TRUE,cpo=TRUE))
  round(model2.par$summary.fixed[,1:5],3)
  DIC_2<-c(DIC_2, model2.par$dic$dic)
  print(model2.par$dic$dic)
}

#DIC vector
print(DIC_2)
""
[1] 6.681565e+48 1.748291e+03 1.745994e+03 1.745208e+03 1.744214e+03 1.743577e+03
[7] 1.743111e+03 1.743089e+03 1.743577e+03 1.744467e+03 1.746373e+03 1.749216e+03
[13] 1.755190e+03 1.775394e+03 1.095150e+84 1.897523e+69 5.012129e+63 4.404118e+61
[19] 1.392719e+61 7.181979e+60 2.166742e+60
""
#Best DIC value
print(min(DIC_2))
""
1743.089
""
#Best DIC model
print(which.min(DIC_2))
#8 p=0.7
""

```

```

'''
###
#Best Model for Deaths has p=0.7
###

p=0.7
formula2.par <- deaths ~ 1 + f(data21$county,model="bym",graph=Ohio.adj, constr=TRUE)
+ f(county1,time,model="iid", constr=TRUE)+ I(time^p)

model2.par <- inla(formula2.par,family="poisson",data=data21,
  control.predictor=list(compute=TRUE),
  control.compute=list(dic=TRUE,cpo=TRUE))
round(model2.par$summary.fixed[,1:5],3)

'''
      mean  sd 0.025quant 0.5quant 0.975quant
(Intercept) -5.221 0.302  -5.820  -5.218  -4.633
I(time^p)    0.229 0.045   0.136   0.231   0.312
'''

```

Appendix C: Full Outputs

This is the output of the code of the Linear BYM model for the number of cases. In particular, these are the estimates of the model parameters b_0 and β in the case $p = 1$ and the value of the corresponding DIC.

```

'''
      mean  sd 0.025quant 0.5quant 0.975quant
(Intercept) -1.030 0.144  -1.316  -1.029  -0.750
I(time^1)    0.124 0.002   0.121   0.124   0.128
'''

```

```

model.par$dic$dic
#11670.49

```

These are the estimates of the posterior means of the spatial main effects ζ_i for $i = 1, \dots, n$.

```

zeta.ST1
index.1 index.2 index.3 index.4 index.5 index.6 index.7
1.52291472 1.15583188 1.14840846 1.17135565 2.63916662 0.69335907 1.64888552
index.8 index.9 index.10 index.11 index.12 index.13 index.14
0.79928097 8.82974544 1.89896521 1.52203934 0.89634585 2.15965239 0.67689111
index.15 index.16 index.17 index.18 index.19 index.20 index.21
0.58980138 2.05710749 0.82884484 14.74848220 0.26663371 1.92479386 3.15618307
index.22 index.23 index.24 index.25 index.26 index.27 index.28
1.29510973 2.87710079 0.20754939 6.15472617 1.21689774 0.72094574 0.88253530
index.29 index.30 index.31 index.32 index.33 index.34 index.35
0.78654809 0.28515287 6.24271809 0.63224499 0.16963868 0.38541384 0.41793440
index.36 index.37 index.38 index.39 index.40 index.41 index.42
0.63235966 0.17614517 0.84153329 0.89137506 1.35372414 3.17126771 1.18042406
index.43 index.44 index.45 index.46 index.47 index.48 index.49
2.81869440 0.34226159 1.66147577 1.44112744 6.55112042 1.17165036 1.30716582
index.50 index.51 index.52 index.53 index.54 index.55 index.56
6.25853211 0.66681800 3.94698431 0.78141386 0.82790099 6.24264254 2.21892673

```

```

index.57 index.58 index.59 index.60 index.61 index.62 index.63
1.46439827 0.82295745 0.44242395 0.98066729 0.65878224 0.70299420 0.79183798
index.64 index.65 index.66 index.67 index.68 index.69 index.70
1.24382218 0.58334702 1.16801973 3.28216210 0.48220495 0.01232118 0.73132563
index.71 index.72 index.73 index.74 index.75 index.76 index.77
0.40069469 0.90647997 1.00178060 1.01962772 1.46198862 2.27714820 4.95697180
index.78 index.79 index.80 index.81 index.82 index.83 index.84
0.72521309 0.64772456 1.15670576 54.92638427 1.47814277 0.54349093 3.14448274
index.85 index.86 index.87 index.88
0.16448801 1.32235705 0.14970093 1.70330616

```

These are the estimates of the time differential effect δ_i for $i = 1, \dots, n$.

```

m
[1] -0.0736230841 0.0070411164 -0.0428609154 0.0043942070 -0.0806315580 -0.0061415887
[7] 0.0024579641 -0.0344092627 -0.0125369768 -0.0341237599 -0.0453234795 0.0029827949
[13] -0.0020813590 0.0055735883 0.0714425221 -0.0271516181 -0.0035139379 0.1033839040
[19] 0.0944549740 -0.0325112268 0.0311765349 -0.0182355138 0.0090516368 0.0163063748
[25] 0.1009295310 -0.0397258546 -0.0278414943 0.0647899043 0.0172910386 -0.0059139225
[31] 0.0577191909 0.0183178075 0.0141448335 -0.0529573543 -0.0478536368 -0.0167793205
[37] -0.0132193427 -0.0552907730 -0.0058364386 -0.0719950595 -0.0337888070 -0.0244351223
[43] 0.0427817194 0.0235198609 0.0364418338 -0.0417927361 0.0359712511 0.1325914545
[49] 0.0001961628 0.0562784113 0.0403980205 0.0379351703 -0.0721675341 -0.0091257893
[55] 0.0115548638 -0.0928527169 0.0660688881 -0.0615967862 -0.0035593920 -0.0291333580
[61] -0.0566115600 -0.0117654571 -0.0479960002 -0.0435058422 0.0466197030 -0.0855581523
[67] 0.0268410734 0.0019266635 0.0775019857 0.0290847197 0.0140334173 -0.0191253316
[73] -0.0547981480 -0.0258499437 -0.0102877317 0.0603662508 0.0672975023 0.1146213580
[79] 0.0480804586 -0.0178295348 -0.0989958891 -0.0761436479 -0.0563006600 0.0045749866
[85] 0.0751317133 0.0130560906 0.0018759080 0.0159569297

```

This is the output of the code of the Linear BYM model for the number of deaths. In particular, these are the estimates of the model parameters b_0 and β in the case $p = 1$ and the value of the corresponding DIC.

```

""
      mean sd 0.025quant 0.5quant 0.975quant
(Intercept) -4.550 0.210 -4.968 -4.547 -4.145
time      0.054 0.013 0.028 0.055 0.078
""

```

```

model2.par$dic$dic
#1746.373

```

These are the estimates of the posterior means of the spatial main effects ζ_i for $i = 1, \dots, n$.

```

index.1 index.2 index.3 index.4 index.5 index.6 index.7 index.8 index.9
1.0035486 0.9968942 0.9989173 0.9998835 1.0060013 0.9980919 0.9991305 1.0053879 0.9990498
index.10 index.11 index.12 index.13 index.14 index.15 index.16 index.17 index.18
0.9990749 0.9983781 0.9989874 1.0022046 1.0005633 1.0032919 1.0035163 1.0019193 1.0008599
index.19 index.20 index.21 index.22 index.23 index.24 index.25 index.26 index.27
0.9967121 0.9990840 1.0035728 1.0098401 0.9989840 0.9989154 1.0040492 1.0010679 1.0074702
index.28 index.29 index.30 index.31 index.32 index.33 index.34 index.35 index.36
0.9989606 0.9999518 0.9990761 0.9935888 0.9987185 0.9990972 0.9992246 0.9997577 1.0002665
index.37 index.38 index.39 index.40 index.41 index.42 index.43 index.44 index.45

```

0.9984995 1.0037205 1.0039863 0.9999408 1.0009392 1.0039843 1.0033154 0.9988733 0.9979998
index.46 index.47 index.48 index.49 index.50 index.51 index.52 index.53 index.54
 1.0029753 1.0028280 1.0087520 1.0025086 1.0034441 1.0030884 1.0032302 0.9999123 1.0007798
index.55 index.56 index.57 index.58 index.59 index.60 index.61 index.62 index.63
 1.0079712 1.0002209 0.9999938 1.0000692 1.0003810 0.9995006 1.0028350 1.0000269 1.0005631
index.64 index.65 index.66 index.67 index.68 index.69 index.70 index.71 index.72
 1.0004931 0.9988251 1.0012808 0.9976089 1.0023559 0.9986939 0.9992733 0.9998042 0.9984645
index.73 index.74 index.75 index.76 index.77 index.78 index.79 index.80 index.81
 1.0004424 1.0034499 0.9992026 1.0051884 0.9997734 1.0012931 0.9992502 0.9989074 1.0022338
index.82 index.83 index.84 index.85 index.86 index.87 index.88
 0.9994354 1.0033337 0.9999106 1.0018419 0.9976054 0.9987744 0.9974478

These are the estimates of the time differential effect δ_i for $i = 1, \dots, n$.

[1] -0.0716048256 0.1059521402 -0.0606503830 -0.0630040940 0.0890923911 0.0658620728
 [7] -0.0544374253 0.0767273593 0.0463886086 -0.0649060815 0.0480425316 -0.0606522325
 [13] 0.0931658401 0.0133081552 0.1629430087 -0.0594382059 -0.0665706717 0.2476725114
 [19] 0.1501242801 -0.0639894867 0.1149590144 0.1185976435 0.0115475733 -0.0671872548
 [25] 0.2057952917 -0.0684423501 0.1110119226 -0.0578347769 0.0999185806 -0.0709316501
 [31] 0.1721012297 0.0414307576 -0.0709321696 -0.0733793571 -0.0724231609 -0.0665301946
 [37] 0.0317070978 -0.0699382551 0.1149431107 -0.0724265837 -0.0678658540 0.0922368384
 [43] 0.1613308635 -0.0671871917 0.1237462020 -0.0658381008 0.1593889357 0.1992120282
 [49] 0.1212705051 0.2203218984 -0.0630885366 0.1348163639 -0.0722314793 0.0830050724
 [55] 0.2044096523 -0.0717925647 0.1334999585 -0.0720217435 -0.0706262001 -0.0678313440
 [61] -0.0726573345 -0.0689415971 -0.0720313742 -0.0702644701 -0.0671848742 -0.0689710418
 [67] 0.1668434873 0.0638022342 0.0589833647 -0.0618796404 -0.0698510467 0.0807199708
 [73] -0.0720289214 0.0907502381 -0.0683986255 0.2129024512 0.2208535735 0.1890151488
 [79] -0.0544410842 -0.0629789660 0.0007678136 -0.0689284827 -0.0741474516 -0.0630041718
 [85] 0.0516243518 0.1133611344 0.0472618516 0.0985899912

-End of Code-