

The Mathematics of COVID-19: An Introduction to Data Driven Analysis for COVID-19 Modeling, Topological Weighted Centroid COVID-19 Summer Program

Masoud Asadi-Zeydabadi¹, Weldon A. Lodwick², and Francis Newman^{2,3} with help from Dr. Massimo Buscema (Semeion Research Center, Rome, Italy) and Dr. Marina T. Mizukoshi (Federal University of Goiás, Brasil)

1. University of Colorado Denver - Department of Physics 2. University of Colorado Denver- Department of Math/Stat, 3. Professor Emeritus Anschutz Medical Campus, Department of Radiation Oncology

July 27, 2020

- 1 The Problem
- 2 Two Approaches
 - 1 Model Driven Analysis Differential Equations - Susceptible, Infected, Recovered (SIR) Model
 - 2 Data Driven Analysis - Topological Weighted Centroid (TWC)
- 3 Susceptible, Infected, Recovered (SIR)
 - 1 Differential Equations
 - 2 Application of the SIR to the COVID-19 Outbreak
- 4 Topological Weighted Centroid (TWC) Details
- 5 TWC Applications
 - 1 USA
 - 2 Brazil

The Problem

- Given data of COVID cases on various dates, determine, for a particular date, the places where the impetus for the spread is most accelerating. That is, from where (what geographical locations) is the spread potentially most “effective” in causing the spread that has produced the collected data?
- Where will the effective location be with respect to causing tomorrow’s (future) spread? That is, given past data on various dates, determine where the new center of accelerated spread will be.
- Determine the likely number of infections and the time course of the disease in order to inform public health officials and health care system responses.

Two Approaches

- 1 Model Driven Analysis - A model that encodes the understanding of the behavior of a system, coronavirus causing COVID-19 in the case at hand, by fitting a predetermined cause/effect relationship to the data. That is, the equations and mathematical relationships (their cause/effect) are the explanation of the what occurs. The data fits into the model. SIR differential equation models are a type of model driven analysis.
- 2 Data Driven Analysis - The model (the set of cause/effect relationships) is elicited from the data.

Introduction - The Mathematics of COVID-19

The World Health Organization (WHO) in China, reported, in December of 2019, a pneumonia with unknown cause in Wuhan, China. The new virus, the coronavirus or SARS-CoV2, caused COVID-19 disease, which spread and is spreading over the world. The first case of COVID-19 in USA was confirmed on January 20, 2020. The Center for Disease Control and Prevention (CDC), on February 26, 2020 confirmed the first case of community spread of COVID-19 in California. The first death from COVID-19 in USA was in Washington state, on February 20, 2020. The number of cases and deaths were accelerating (March 26, 2020, according to Johns Hopkins University, Coronavirus Resource Center). As of March 26, 2020, the number of confirmed and recovered cases and death are 85,991, 753 and 1,296 respectively.

Introduction

- This talk presents both a model driven differential equation model called the SIR epidemic model and a data driven model called the Topological Weighted Centroid (TWC) algorithms applied to the analysis of COVID-19.
- Both of these mathematical tools have been successfully used for epidemic analysis. The CDC uses, among other things, a more complex version of SIR. Italy has used the TWC algorithms to predict the pattern and movement of the COVID-19 outbreak.
- In addition Semeion Research Center, Rome, Italy (the research center that developed TWC) has used TWC to not only look at the geographic pattern of epidemics, but the origin of a variety of disease outbreak around the world.
- UCD's Center for Computational and Mathematical Biology and Semeion collaborate. In fact, the Italian and USA state departments officially designated our centers as mutual collaborators.

Compartment Models

Prior to looking at TWC, a data driven system, applied to COVID-19, let us look a non-statistical mathematical way of encoding cause/effect relationships called compartmental models, where equations are used to encode the cause/effect relationships. Mechanistic models, in this case compartment models, have been around in the literature since at least 1927. They offer important features.

- 1 The compartments in the case of COVID-19 vary according to the complexity of the model . One simple version has the following compartments: *susceptible*, *infected*, *removed* or SIR. Each compartment is represented by an equation with links to other compartments that are understood to be some of their common causes/effects.
- 2 The encoded hypothesized causal characteristics of disease transmission makes it easier to find points of intervention.

Susceptible/Infected/Removed (SIR)

S: Susceptible \rightarrow I: Infected \rightarrow R: Removed (see

<https://www.youtube.com/watch?v=NKMHHm2Zbkw> for COVID-19).

The SIR model hypothesizes that

- 1 The rate of change at which the susceptible people are removed from their compartment is proportional to the encounter of the susceptible population with the infected population as a product of the two populations. The constant of proportionality we denote beta β . That is, The **rate of change in $S = -\beta IS$** .
- 2 The rate at which the populations is removed (become immune or die) is proportional to the infected population. That is, the **rate of change in $R = \gamma I$** . The parameter γ is the mortality/recovery (removal) rate.
- 3 The rate at which the infected population changes depends on the rate at which susceptible population changes (from susceptible to infected) minus the rate at which infected are removed (become immune or die). That is, the **rate of change in $I = \beta IS - \gamma I$** .

Parameters of the SIR Model

- 1 The average period associated with the contraction of the coronavirus from the contact with an infected population is denoted B and is equal to $\frac{1}{\beta}$ where β is the constant of proportionality of the relation between susceptible and infected. That is $B = \frac{1}{\beta}$.
- 2 The average period of infection is denoted C and is equal to $\frac{1}{\gamma}$, where γ is the constant of proportionality of the relationship of removal to infected. That is, $C = \frac{1}{\gamma}$. the time during which an individual can infect others.
- 3 The basic reproduction number is the product of the contact rate and the infectious period $R_0 = \beta C = C/B$, delineating the number of individuals that are infected by one individual in an otherwise uninfected, susceptible population. The basic reproduction number R_0 can be thought of as the number of secondary cases which one case would produce in a completely susceptible population where $R_0 > 1$ implies outbreak and $R_0 < 1$ implies attenuation of outbreak. The following simulation uses data for USA, where $R_0 = 5.8$.

Simulation SIR with USA parameters

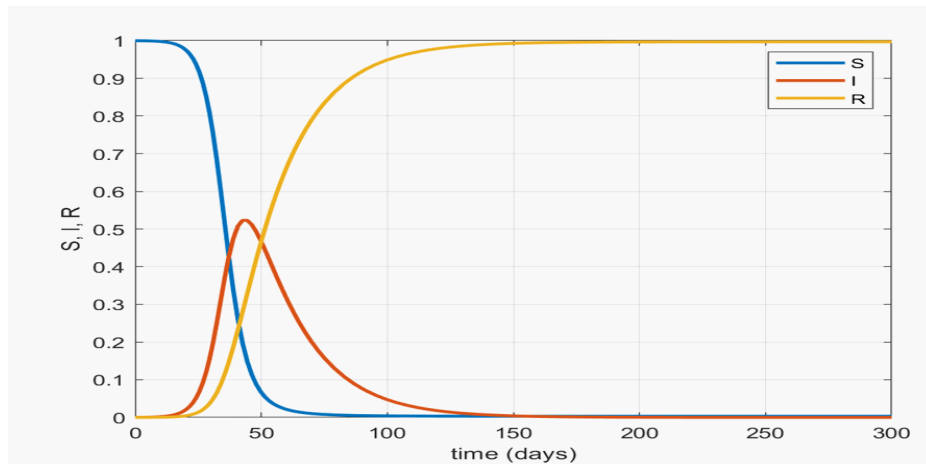


Figure: SIR Model $\beta = 0.29$, $\gamma = 0.05$, $S_i = 0.99983$, $I_i = 0.00017$ and $R_i = 0$

- TWC is a data driven geographic dynamic model that consists of six sets of algorithms. *Geographic* because COVID occurs in space. *Dynamic* because COVID “spreads” in time and there is a rate (speed) of spread associated with COVID.
- The idea behind TWC is best explained by an analogy. TWC is the mathematization of a minimal number of “coronavirus terrorist cells”, whose aim is to be located in spots across the geographical region (March 2020 USA for example) in time so that they might most efficiently actualize the distribution of cases as indicated by the time-changing data. In other words, if one were a coronavirus terrorist with the aim of having distributed the COVID cases as indicated by the data at particular time, where would be the location from which the actual distribution of cases would most efficiently have emanated?

- Another way of saying this is: If the COVID cases each represented terrorist soft targets, where would a terrorist cell locate itself to have the greatest chance of actualizing damage utilizing the optimum energy for the given “terrain” .
- The mathematization of this is achieved via statistical thermodynamics, optimization of free energy, and entropy.
- What TWC constructs, under various types of points of view, for example, is the location of where a terrorist would locate itself, have its operational center, to maximize the chance of destruction given the configuration of soft targets. Another point of view is where a predator would locate itself in anticipation of a future configuration of prey.

We will focus on just three of the eighteen different algorithms that considers the number of cases per location.

- 1 TWC- α map locates the geographical area where a coronavirus terrorist would have located itself yesterday in order to have achieved the distribution of cases of the disease today.
- 2 TWC- β map locates the geographical area where a coronavirus terrorist would locate itself today in order to achieve the distribution of cases of the disease tomorrow.
- 3 TWC- γ map predicts of the next steps of the terrorist cells' location to achieve future epidemic spread.

TWC-alpha Map

Figure 1 shows the results of The TWC- α for mid-March 2020 data. Two possible outbreak origins are located in west, Washington and California states. There are several isolated hotspots for example we see a hotspot in northeast.

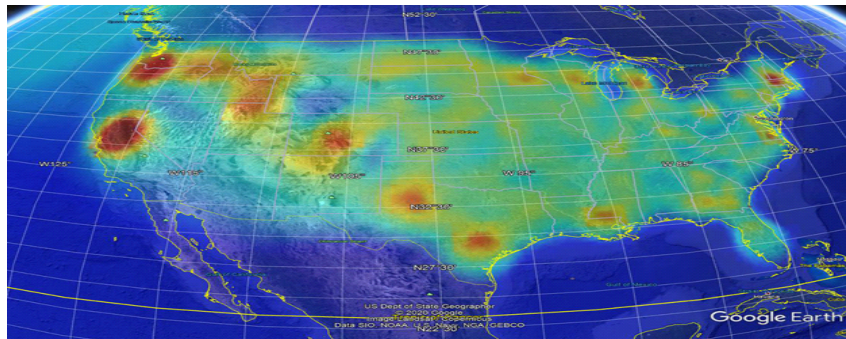


Figure: 1 TWC- α

TWC-beta Map

The result of $TWC-\beta$, Figure 2, shows the outbreak has been expanded over the country (later mid-March). The cold (blue color) places get smaller and turn toward the green and yellow color. From this figure we can see the pattern disease spread over the country.

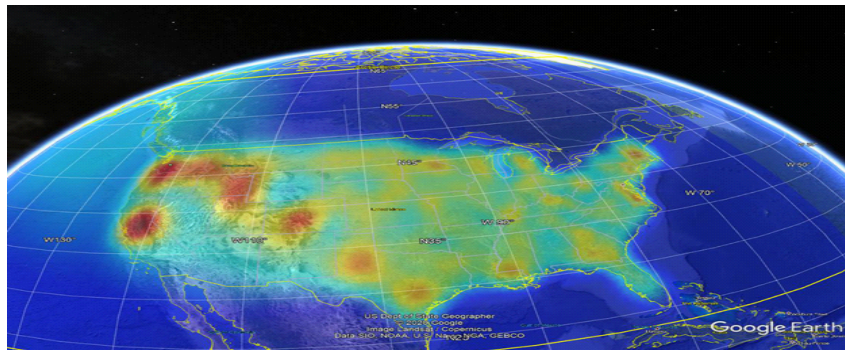


Figure: 2 $TWC-\beta$

TWC-gamma

The result of TWC- γ shows the outbreak in future, Figure 3 later March 2020. This algorithm predicts the recovery in the cold regions is quicker and those places do not become the center of the *COVID-19* outbreak.

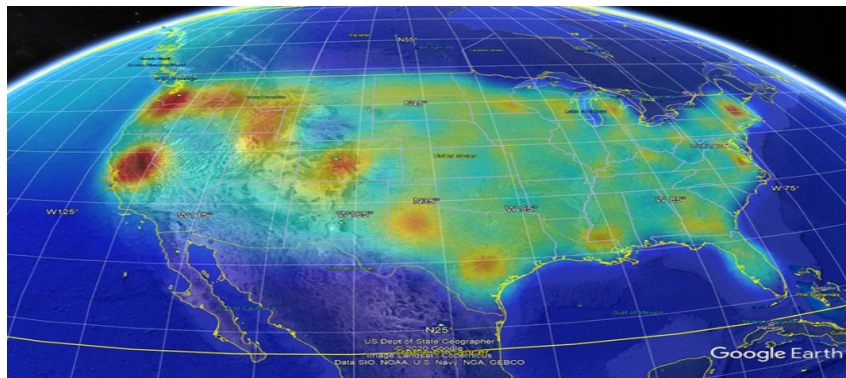
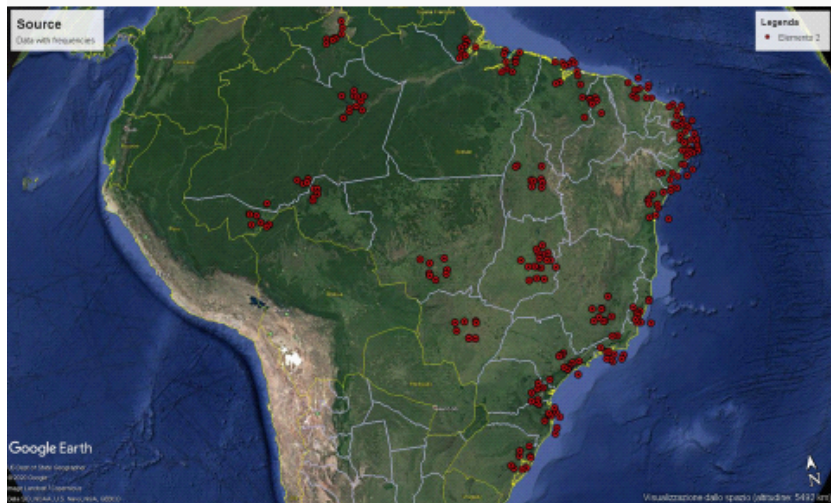


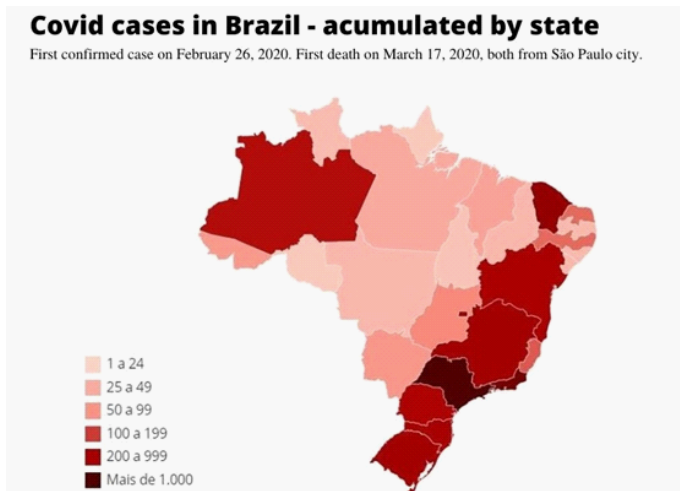
Figure: 3:TWC- γ

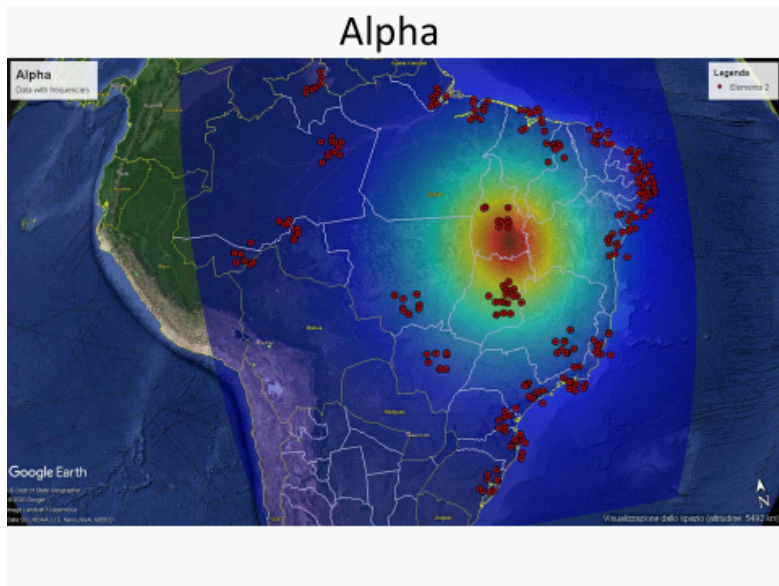
Source Data



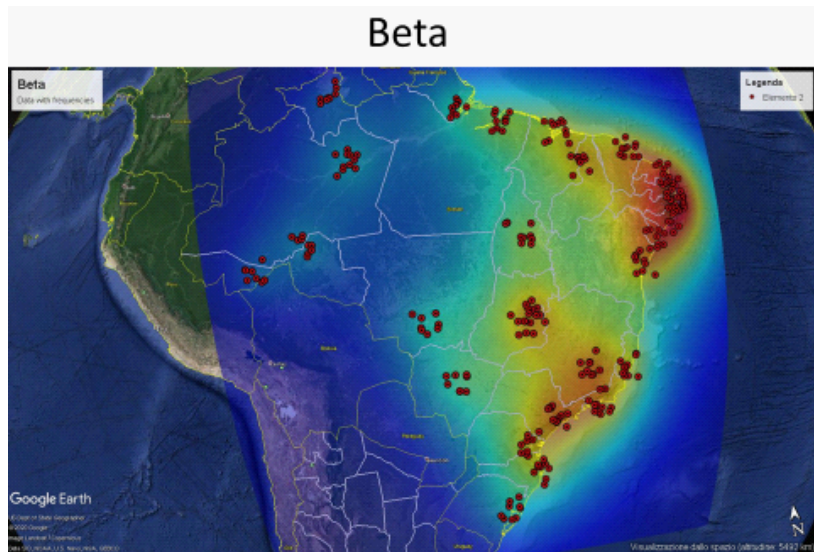
Confirmed COVID Cases in Brazil Through March 17, 2020

Where would the coronavirus "terrorist" locate itself to affect the distribution on the previous slide and summarized by this slide? TWC theory summarizes this on the TWC- α map of the next slide.



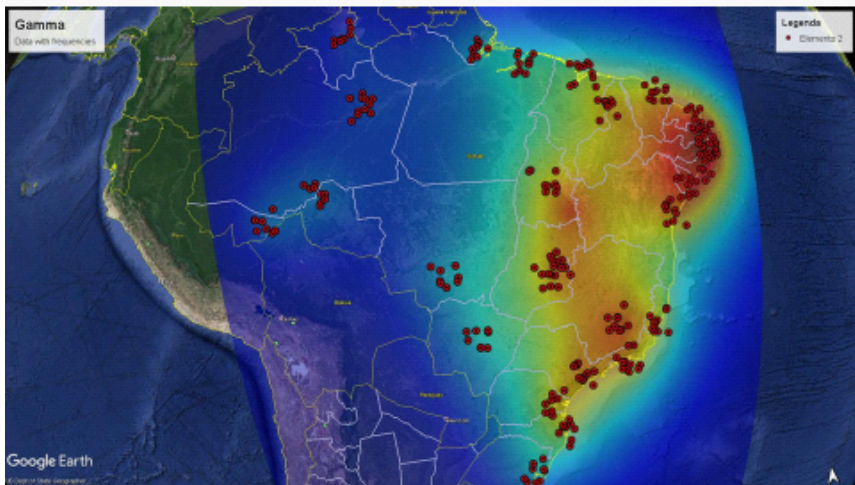


Affect tomorrow's distribution in Brazil.

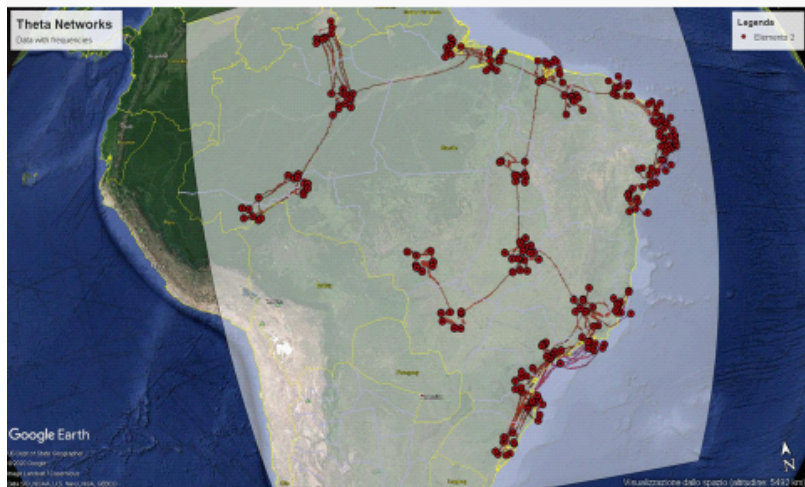


The future impetus for the spread of the coronavirus in the future.

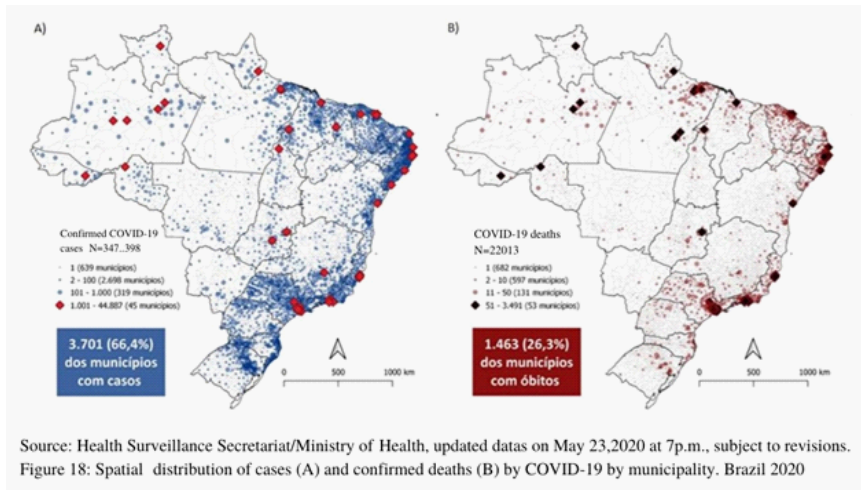
Gamma



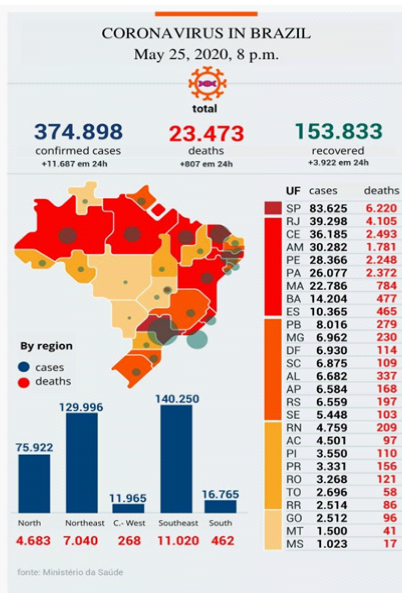
Theta Networks



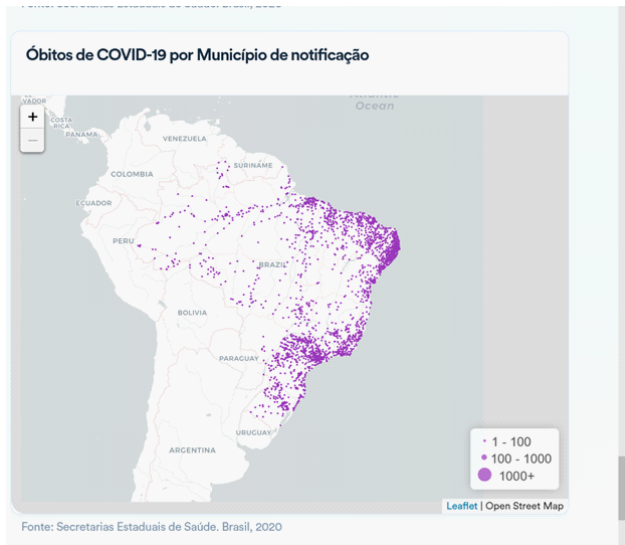
Brazil data 1 - COVID May 23, 2020



Brazil data 2



Brazil data 3 - Municipality Cases June 2, 2020



Selected Bibliography - Web Resources

- <https://www.who.int/emergencies/diseases/novel-coronavirus-2019/events-as-they-happen>
- https://wwwnc.cdc.gov/eid/article/26/6/20-0233_article
- <https://www.cdc.gov/media/releases/2020/s0226-Covid-19-spread.html>
- <https://coronavirus.jhu.edu/map.html>
- https://github.com/CSSEGISandData/COVID-19/blob/master/csse_covid_19_data/csse_covid_19_daily_reports/03-26-2020.csv
- https://github.com/CSSEGISandData/COVID-19/blob/master/csse_covid_19_data/csse_covid_19_daily_reports/02-20-2020.csv

Selected Bibliography - Publications

- Buscema M., Massini G. and Sacco P. “The Topological Weighted Centroid (TWC): A topological approach to the time-space structure of epidemic and pseudo-epidemic processes”, *Physica A*, (2017).
- Buscema, M., Grossi, E., Bronstein, A., Lodwick, W., Asadi-Zeydabadi, M., Benzi, R. & Newman, F. “A new algorithm for identifying possible epidemic sources with application to the German Escherichia Coli outbreak”, *ISPRS Int. J. Geo- Inf.*, **2**, pp. 155-200, (2013); DOI:10.3390/ijgi2010155
- Massimo Buscema, Masoud Asadi-Zeydabadi, Weldon Lodwick, Alphonse Nde Nembot, Alvin Bronstein and Francis Newman, “Analysis of the Ebola Outbreak in 2014 and 2018 in West Africa and Congo by Using Artificial Adaptive System”, *Applied Artificial Intelligence*, 2020

Selected Bibliography - Publications

- Holshue ML, DeBolt C, Lindquist S, Lofy KH, et al. (March 2020). “First Case of 2019 Novel Coronavirus in the United States”. *N. Engl. J. Med.* **382**:10, pp. 929–936. doi:10.1056/NEJMoa2001191
- Mathias Peirlinck, Kevin Linka, Francisco Sahli Costabal, and Ellen Kuhl, “Outbreak dynamics of COVID-19 in China and the United States”, *Biomech Model Mechanobiol.* 2020 Apr 27, pp. 1–15.