Chapter 6

COVID-19 in South America-Why Geometrical Risks Must be combined with Statistical Analysis to Improve Emergency Response

Gauss M Cordeiro[1,8]*; Jones O Albuquerque[2,3,8]; Fernando Nóbrega[2,4,8]; Danillo Souza[2,7,8]; Allisson Dantas[2,8]; Daniel Lopez-Codina[6,8]; José Luiz de Lima Filho[2,8]

[1]*Department of Statistics, Universidade Federal de Pernambuco - UFPE,*

[2]*Keizo-Asami Immunopathology Laboratory, UFPE*

[3]*Department of Statistics and Informatics, Universidade Federal Rural de Pernambuco - UFRPE,*

[4]*Department of Mathematics, UFPE,*

[5]*Institute of Advanced Studies, University of Amsterdam,*

[6]*Department of Physics, Universitat Politécnica de Catalunya,*

[7]*Basque Centre for Applied Mathematics - BCAM, Bilbao, 8Institute for Risk and Disaster Reduction of Pernambuco, UFRPE.*

[8]*Institute for Risk and Disaster Reduction of Pernambuco, UFRPE.*

[*]***Correspondence to: Gauss M Cordeiro,*** *Department of Statistics, Universidade Federal de Pernambuco – UFPE.*

Email: *gauss@de.ufpe.br*

1. Introduction

The new coronavirus was discovered in the Chinese province of Wuhan in late December 2019 and was named SARS-CoV 2 due to its similarity with SARS-CoV. Infections caused by this agent are now called COVID-19.

Some of the most influential global websites on Coronavirus are "Worldometer" [1], "Our World in Data" [2], "Johns Hopkins Research Center" [3], and "World Health Organization (WHO) Coronavirus Dashboard" [4].

Worldometer [1] aggregates data from thousands of sources in real time and provides updated global COVID-19 statistics from 220 countries and territories such as total and new

Citation: Cordeiro GM , (2021) Latest updates on SARS-CoV-2 (Corona Virus), Vol. 1, Chapter 6, pp. 48-56.

cases,total and new deaths, total and new recoveries, active and critical cases, total cases and total deaths per million, total tests, tests per million and population. Plots for the new daily cases and deaths and 3-day and 7-day moving averages are reported for each country.

The user of Our World in Data [2] can select the countries and obtain plots for the daily new confirmed cases and deaths per million people, estimates of the effective reproduction rate of COVID-19, numbers of tests, tests conducted per confirmed case, case fatality rate, people who received at least one dose of vaccine, people fully vaccinated, hospital admissions, intensive care unit admissions, among other plots.

Johns Hopkins [3] reports for every country (entire time, past day, past week and past month) the confirmed cases and deaths, vaccine tracking, doses administered, people fully vaccinated and percentage of population fully vaccinated. It provides more than 100 plots for selecting and comparing countries and regions. These include plots of confirmed cases and deaths and rolling 7-day averages, confirmed cases and deaths per million people and rolling 7-day averages, cumulative confirmed cases, cumulative confirmed cases and deaths by region, daily new confirmed cases *vs.* cumulative cases, cumulative confirmed cases per million *vs.* gross domestic product (GDP) per capita, confirmed deaths per million *vs.* GDP per capita, confirmed deaths *vs.* population density, cumulative confirmed deaths *vs.* cases, death rate *vs.* population density, map and country time-series, vaccine doses administered, vaccine doses administered by manufacturer, vaccine doses administered per one hundred people, variants in analyzed sequences, containment and health index, daily tests *vs.* daily new confirmed cases, daily tests *vs.* daily new confirmed cases per million, stringency index, case fatality rate *vs.* median age of the population, case fatality rate of the ongoing pandemic, case fatality rate *vs.* total confirmed deaths, daily tests per thousand people, daily tests per thousand people and rolling 7-day average, daily vaccine doses administered, daily vaccine doses administered per hundred people, among several others.

Finally, the WHO Coronavirus Dashboard [4] reports (by region, country, territory and area) the cumulative total of cases and deaths and newly reported cases and deaths in the last 24 hours,and plots (daily and weekly) of confirmed cases and deaths by region and country.

Using all the functionalities of these websites seems insufficient to provide governments with the future scenarios of the dynamic behavior of COVID-19 in the world. These difficulties are due to several reasons. The diagnostic gold standard method to detect the presence of the virus is the LAMP-PCR, or RT-PCR. The former method takes only a few minutes but can only be used for a few samples at once. For large numbers of samples, RT-PCR is still the best, but it takes several hours. Both are too expensive for many countries. Based on that, the number of cases detected daily is underestimated. The numbers of infected people rise daily in different countries due to human behaviour, lack of vaccination, and new virus variants. These usually happen in countries with large population and substantial numbers of infected people, such as Brazil and India, giving the virus chance to mutate into a more contagious variant. And for viral detection, samples should be sequenced in a specific laboratory with well-trained personnel, which is very difficult to achieve in many countries. These difficulties make the control of pandemics difficult. To help understand this pandemic, it is necessary to use models that do not need only numbers based on molecular testing but also based on virus sequencing [19].

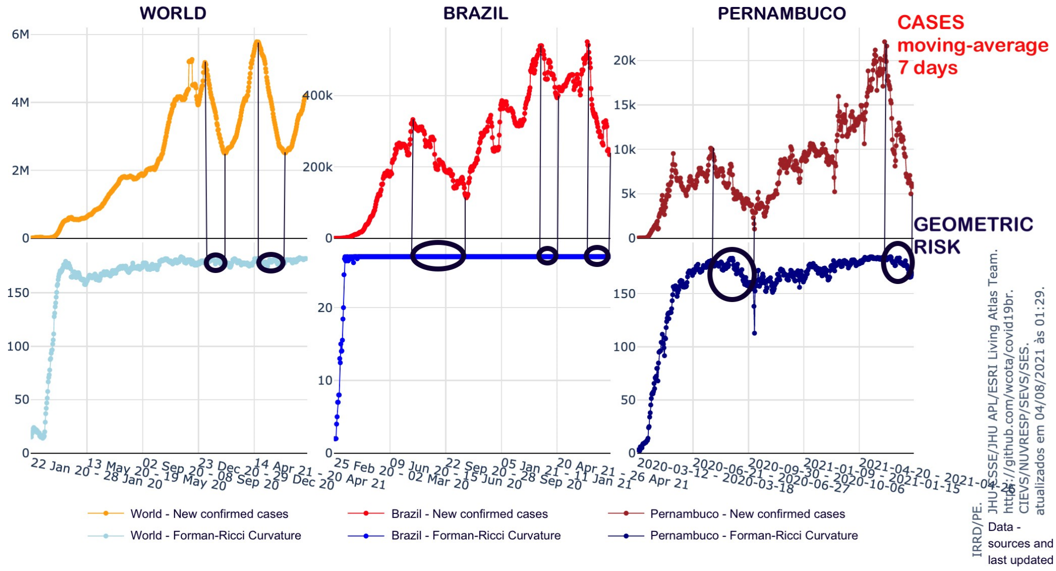

Figure 1: New cases versus Ricci curvatures (non-filtered version) [18], which we call a geometric risk approach. 7-day moving average showing down cases periods amid high pandemic risk.

In order to improve the accuracy of analyses and their scenarios, several scientists have applied dynamic approaches based on mathematical epidemiological models [16]. Compartmental models, $R(t)$ metrics, and other previous tools have been used to capture the dynamic behavior of COVID-19, but several of them have been found to be wrong or misunderstood in their use [15, 17].

Some misconstrued or even wrong uses of these metrics and approaches prompt policymakers to formulate misguided public policies, such as early relaxation of social distancing, etc., and cause the public to lose the situational awareness of risks to life, as illustrated in **Figure 1** obtained from the "IRRD" [5]. In this article [18], the authors presented a geometric, data-driven parameter free approach to COVID-19 analysis, to show the use of the Ricci curvatures as a pandemic indicator. Figure 1 illustrates the geometry of the worldwide new cases in contrast with Brazilian and Pernambuco (Brazilian state) data. The values of Ricci curvature can be interpreted as the level of connections between cases (we call approaches like this, based on geometry, **GEOMETRIC RISK**). This pattern remains over time, even when the new case curves appeared to decay. In this chapter, we present some geometric data supported by solid statistical analyses to show how seriously the pandemic is in a place where stochastic and statistical data alone does not always produce reliable insights.

2. Covid-19 in South America

South America was one of the last continents to confirm an infection by the new coronavirus. The pandemic came to Brazil from the arrival of foreigners and the return of Brazilians from Europe at the country's main international ports and airports. The first confirmed case in Brazilwas on February 26, 2020, although researchers have described indications of cases as early as January. The first cases in the other South American countries were confirmed during March. The governments of these countries have adopted different strategies to try to reduce the transmission of the virus, increase health infrastructure, and financially support workers and companies.

The WHO began to warn that South America was a new epicenter when the cases began to increase significantly between April 7-13, 2021 in several countries as can be noted at the websites [1-3]. Indeed, in this period, the 7-day moving averages (MAs) of confirmed cases increased in Argentina, Peru, Chile and Uruguay by approximately 49%, 40%,38% and 22%, respectively.

South America has only 5.5% of the world's population, but one in four deaths from coronavirus in the world has been recorded in the continent, which leads to the highest mortality rate among continents with about 2,573 deaths per million residents (on July 31, 2021) [1]. The population of Europe is 1.74 times larger than that one of South America. However, the total number of COVID-19 deaths is approximately one million in both continents [1-3].

In Brazil, the largest nation in South America, the scale of the pandemic has also taken on new dimensions, with almost 20 million confirmed cases and more than 556 thousand deaths according to [7]. Brazil is the third country in the world in number of cases, behind the UnitedStates and India, and it is second in terms of deaths behind the United States [1-3]. However, Brazil may reach the first place of deaths in the very near future if the public policies continue to be based solely on numbers and moving averages instead of including pandemic risk metric strategies.

The pandemic has caused the biggest public health crisis in South America of all time. The most vulnerable people have suffered the most since the start of the pandemic because the public healthcare system has become overwhelmed in several cities due to the increasing numbers of coronavirus cases. The hospital beds, health workers and general medical supplies were in short supply long before the start of the pandemic. The structural problems caused in the region will take a long time to be solve. One of the main reason for this tragedy is that South America has the greatest structural inequality in the world. In fact, eight countries in South America are among the 20 most economically unequal countries in the world. About 30% of the South American population lives below the poverty line and one in five people live in shanty towns or other irregular settlements. Nearly 75% of residents in this region have low or lower-middle income and only 3% of them are classified as having high income [8]. The six most unequal countries in South America based on the Gini coefficient (a higher Gini index indicates greater inequality) were (in 2017): Brazil (53.3), Colombia (49.7), Paraguay (48.8), Chile (46.6), Ecuador (44.7), and Bolivia (44) [9]. Political leaders from many countries took a long time to adopt policies of isolation or even social distancing and the mandatory use of masks in closed spaces. A large contingent of people live in extreme poverty and have to ignore or violate these policies to survive. More than 50% of the South American workers are in the informal sector and have to move around to perform services or sell goods. Very poor people literally live from hand to mouth. The flawed pandemic emergency response and belated actions in South America can be observed by a geometric risk diagram. Some South America countries are presented in **Figure 2**. For example, in Brazil, Ecuador and Peru, the misdirected actions can be observed by the multiple trajectories of the daily calculated COVID-19 risk in the red zone of the risk diagram. All the South America risk diagrams can be automatically obtained daily at https://irrd.tech/interactive-risk-diagrams/southamerica/. Linear (not exponential or logarithmic) and straight (no sudden changes in their direction) trajectories as in Chile, Colombia and Uruguay, from red zone to green zone indicate well-defined public strategies.

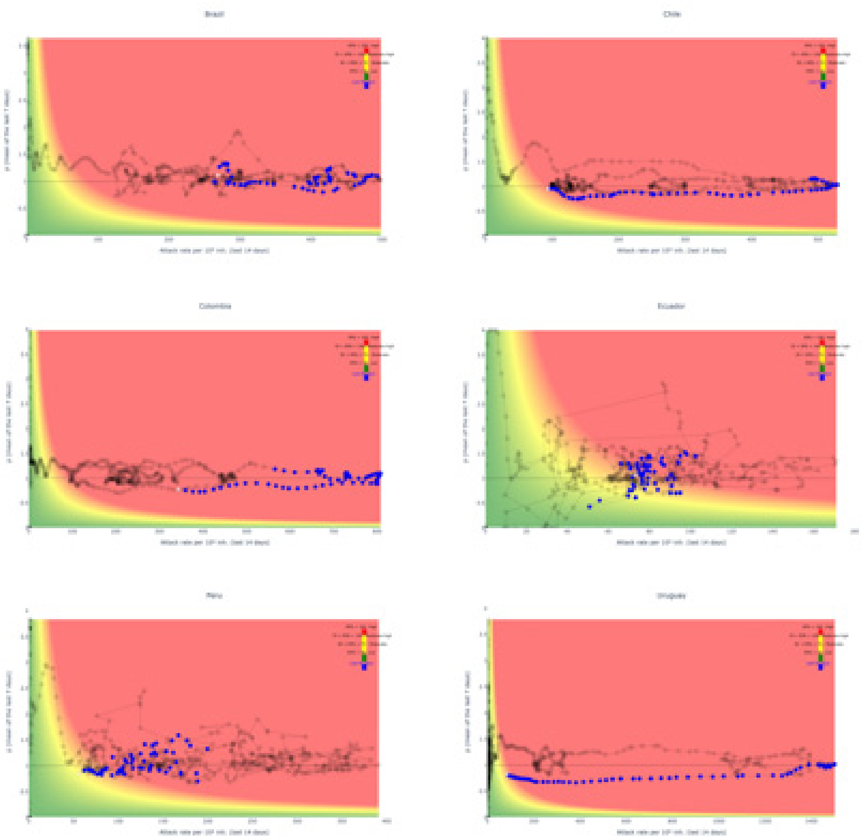

Figure 2: For each country, white point-last estimated potential growth (EPG) day, blue points- last 30 day EPG, black points-past EPG.

Geometrical Risk Analysis - Emergency Response in South America and Exemplary Countries

The temporal evolution of the COVID emergency responses in countries from different continents and South American countries can be compared using the daily geometric and risk analysis of the countries, respectively. On this approach, the red, orange, yellow and green zones of the risk diagrams mean high, moderate-high, moderate and low risk, according to estimated potential growth (EPG) calculation, which is

$$EPG = \rho_7 \times A_{14},$$

where ρ_7 is the 7-day average propagation velocity (last 3-day number of infections/last 5-day number of infections), and A_{14} is the 14-day accumulated incidence/100,000 people. The explicit formulas are available in the emergency response dashboards for COVID-19 from the "IRRD" and "BIOCOMSC" [*5, 6*]. Table 1 summarizes the EPG calculation ranges, risk levels and colors.

Good strategies to slow transmission are observed on the risk diagrams as linear and straight well-oriented movements of the daily risk (EPG) in direction to the green (safe) zone. For example, Portugal and United Kingdom in Europe, Malawi and South Africa in Africa, and Australia and New Zealand in Oceania have very good past and present movements toward the safe green zone, as illustrated in **Figure 3**.

Table 1: Estimated Potential Growth (EPG), risk level, color.

EPG	risk level	color
EPG ≥ 100	High	Red
70 ≤ EPG < 100	Moderate-high	Orange
30 ≥ EPG < 70	Moderate	Yellow
EPG < 30	Low	Green

3. Basic Statistics, Correlation and Regression

There had been at least 35,538,060 reported infections and 1,090,032 registered deaths caused by the novel coronavirus in South America on July 31, 2021 [1]. The number of COVID-19 cases has now passed one million in five South American countries. This happens for obvious reasons, since Brazil, Argentina, Colombia, Peru and Chile have populations greater than 214, 45, 51, 33 and 19 million people, respectively. The highest numbers of cases are in Brazil, Argentina and Colombia, while the highest death tolls are in Brazil, Peru and Colombia. Uruguay has the highest incidence rate, and Ecuador the lowest. Peru has the highest mortality rate, and Ecuador the lowest. Also Peru's death rate is the highest and Uruguay's the lowest. Among the 20 countries with the highest mortality rates in the world, five are in South

America [1].

The moving averages of confirmed cases and deaths in South American countries can be seen at the websites described in Section 1. For example, we note from [3] that the shapes of the plots of the daily new confirmed COVID-19 cases in Brazil and South America are similar but very different from Argentine and Colombia.

The beta regression pioneered by Ferrari and Cribari-Neto [10] is one the main statistical tools for regression modeling of proportional data. For South American countries, we can consider the demographic, economic and health variables described below as independent variables for constructing regression models to explain part of the variability of their mortality rates. The only work would be to obtain these rates at least at four different times since the beginning of the pandemic based on the method described by Cordeiro et al. [11] for Western European countries. Of course, the different public policies to control the pandemic adopted by countries cannot be included in the regression modeling.

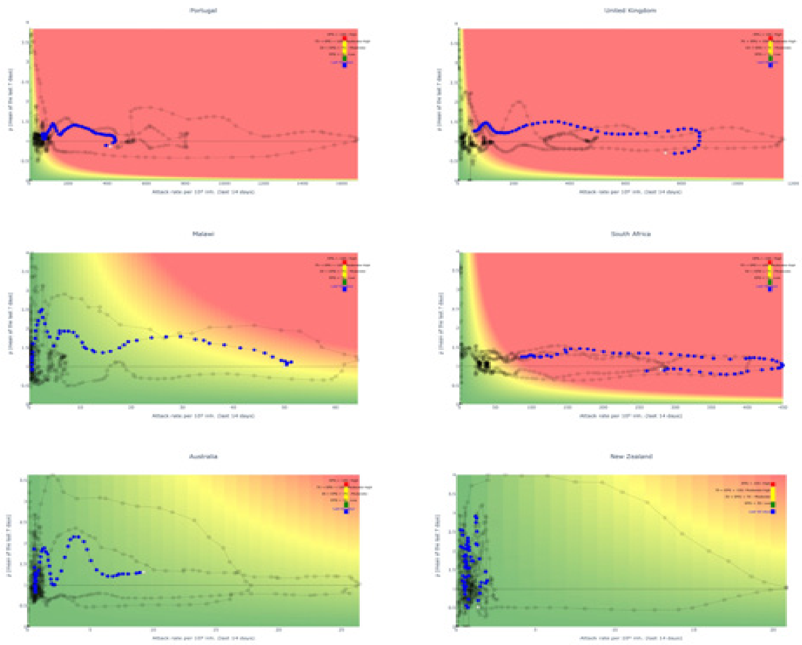

Figure 3: Linear and straight trajectories of the countries' EPG indicate well-defined strategies to combat the pandemic in these countries.

We provide a simple regression to explain part of the variability of the mortality rates of nine South American countries for non-specialist readers. We adopt the following independent variables, whose ranges are obtained from [12]. The population density (number of people per square kilometer) (POPD) varies from 10.6 (Bolivia) to 31.41 (Venezuela); the hospital bed density (beds/1,000 people) (BEDS) ranges from 0.8 (Paraguay) to 2.1 (Brazil and Chile); the physician density (physicians/1,000 people) (PHYD) goes from 1.37 (Paraguay) to 2.44 (Chile); the gross domestic product per capita (GDPPC) based on purchasing power parity (GDP-PPP) (in American dollars) varies from 7,600 (Bolivia) to 24,600 (Chile); the health expenditures (HEXP) (% of GDP) range from 5 (Peru) to 9.5 (Brazil), and the population below the poverty line (in percentage) (POVR) varies from 4.2 (Brazil) to 28 (Colombia). The poverty definition can vary among nations, although the World Bank [13] indicates the global poverty line at US$ 1.90 per day based on the 2011 PPP. The dependent variable is the mortality rate (MR) (on July 18, 2021).

We provide some findings from the correlation matrix of all variables displayed in **Figure 4.** The independent variables BEDS, PHYS, HEXP and GDPPC are all moderately positively correlated. The explanatory variable POVR is moderately negatively correlated with GPDPC, HEXP and PHYS. The mortality rate has moderate negative correlations with HEXP and PHYS but it is uncorrelated with GDPPC. In summary, both hospital bed density and GDPPC have noinfluence on the mortality rate. However, MR decreases when HEXP and PHYS increase.

For the beta regression, we define the response as the MR divided by 10,000, say $y_i = MR_i/10{,}000$, for nine countries (except Venezuela and the territories) and consider HEXP and PHYHD as the explanatory variables. Further, we assume that $y_i \sim Beta(\mu_i, \sigma)$ has a reparameterized beta distribution, where $E(y_i) = \mu_i$ is the expected mortality rate and σ is a dispersion parameter in the variance $Var(y_i) = \sigma^2\mu_i(1 - \mu_i)$ (for i = 1, . . . , 9). Clearly, μ_i and σ are parameters belonging to the (0, 1) interval. The systematic component of the beta regression is

$$\eta_i = \log[\mu_i/(1 - \mu_i)] = \beta_0 + \beta_1 \text{ HEXP}_i + \beta_2 \text{ PHYHD}_i \quad \text{and} \quad \log[\sigma/(1 - \sigma)] = \delta_0.$$

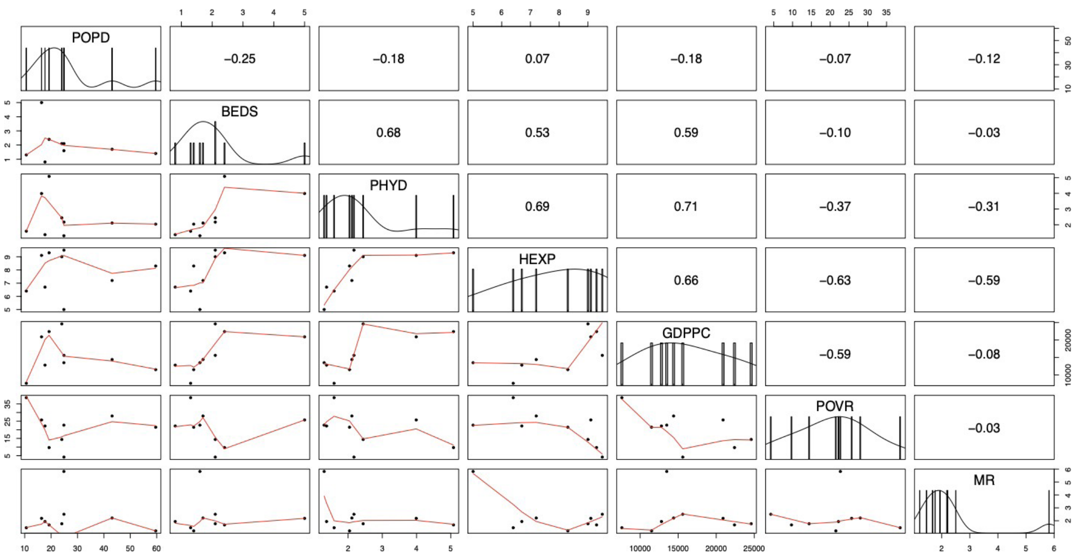

Figure 4: Correlation matrix of the variables.

We fit the beta regression to the current mortality rates by the maximum likelihood method in the **GAMLSS** package of the R software [14]. The explanatory variable PHYHD was not significant, so it was deleted from the above systematic component. The maximum likelihood estimates (MLEs), standard errors (SEs) and *p*-values of the reduced fitted regression are reported in **Table 2**.

Table 2: MLEs, standard errors and *p*-values.

parameter	estimate	SE	p-value
β_0	0.7080	0.9349	0.4776
β_1	-0.2456	0.1227	0.0923
γ_0	-1.2507	0.2846	0.00459

In summary, the proportion of health expenditures explains the variability of the mortality rates in South American countries at a significance level of 0.10. Further, we can note based on the randomized quantile residual plot that the beta regression $\eta_i = 0.7080-0.2456\ \mathrm{HEXP}_i$ is well-fitted to the mortality rates. Thus, these rates will be lower for countries with higher health expenditures in proportion to GPD.

We can also study the effects of the independent variables on the mortality rates marginally. The marginal effects of any independent variable (say x_j) on the response variable MR (assuming the other variables fixed) are obtained by adding one to x_j and calculating the marginal relative difference in the mortality rate.

We can write $\Delta^{(j)}\eta = \beta_j$, and then obtain after some algebra

$$\Delta^{(j)}\mu = \frac{e^{\eta}\, e^{\Delta^{(j)}\eta}}{1 + e^{\eta}\, e^{\Delta^{(j)}\eta}} - \mu.$$

The previous expression can be rewritten as $\Delta^{(j)}\mu/\mu = e^{\beta_j} - 1$. So, assuming that all other factors are fixed, an increase of 1% in the percentage of health expenditure proportion of GDP yields a decrease of 22% in the estimated COVID mortality rate in South America.

4. Concluding Remarks

The emergency response to the COVID-19 pandemic imposes temporalities and demands speedy actions by public policymakers never seen before in the world. The dynamics imposed by SARS-CoV-2 and its variants are challenging researchers in several areas of knowledge to develop better diagnostic technologies and vaccines, as well as requiring public officials to improve education, housing and transport. The statistical and mathematical methods used in epidemiology to monitor and project the severity of the pandemic have also been seriously questioned, since some of them often have indicated misleading scenarios.

Thus, this chapter has presented the numerical-statistical conditions highlighting socio-economic-demographic aspects explaining why several countries in South America are in a very vulnerable situation in the fight against COVID-19. We also have presented how geometric risk analysis methods can help to better capture disease dynamics. Finally, the geometric risk trajectories explained how some South American countries were ineffective in defining and executing strategies to fight the pandemic compared to some countries even in South America and in Europe, Africa and Oceania.

5. Acknowledgments

This work was partially supported by INES (www.ines.org.br), CNPq grant 465614/2014-0, CAPES grant 88887.136410/2017-00, and FACEPE grants APQ-0399-1.03/17 and PRONEX APQ/0388-1.03/14.

6. References

1. COVID Worldometer. Available online: https://www.worldometers. info/coronavirus/#countries (accessed on 31 July 2021).

2. Our World in Data. Available online: https://ourworldindata.org/ (accessed on 31 July 2021).

3. Johns Hopkins Research Center. Available online: https://coronavirus. jhu.edu/map.html (accessed on 31 July 2021).

4. WHO Coronavirus (COVID-19) Dashboard. Available online: https:// covid19.who.int/ (accessed on 31 July 2021).

5. IRRD - Institute for Risk and Disaster Reduction of Pernambuco - Brazil. Universidade Federal Rural de Pernambuco - UFRPE. Available online: https://irrd.org/covid-19 (accessed on 31 July 2021).

6. BIOCOMSC - Computational Biology and Complex Systems. Universitat Polite`cnica de Catalunya.BarcelonaTech. Available online: https:// biocomsc.upc.edu/en/covid-19 (accessed on 31 July 2021).

7. Coronavírus Brasil. Available online: https://covid.saude.gov.br/ (accessed on 31 July 2021).

8. Center for Strategic and International Studies. https://www.csis.org/ analysis/covid-19-exposes-latin-americas-inequality (accessed on 31 July 2021).

9. Statista. https://www.statista.com/statistics/980285/ income-distribution-gini-coefficient-latin-america-cari\-bbean-country/ (accessed on 31 July 2021).

10. Ferrari, S. and Cribari-Neto, F. Beta regression for modelling rates and proportions. Journal of Applied Statistics, 31, 799-815 (2004).

11. Cordeiro, G.M., Ortega, E. M. M., Figueiredo, D., Silva, L, and Prataviera, F. Explaining COVID-19 Mortality Rates in the First Wave in Europe. Model Assisted Statistics and Applications 16, 211-221 (2021).

12. IndexMundi. https://www.indexmundi.com/ (accessed on 31 July 2021).

13. World Bank Group. https://datahelpdesk.worldbank.org/knowledgebase/articles/746163-what-is-the-1-90-poverty-line-and-based-on-this accessed on 31 July 2021).

14. Rigby, R. A. and Stasinopoulos D. M. Generalized additive models for location, scale and shape (with discussion). J. R. Stat. Soc. Ser. C (Appl. Statist.), 54, 507- 554 (2005).

15. Adam, David. A guide to R — the pandemic's misunderstood metric. Nature 583, 346-348 (2020).

16. Adam, David. Special report: The simulations driving the world's response to COVID-19. Nature 580, 316-318 (2020).

17. Inga Holmdahl, S.M., and Caroline Buckee, D.Phil. Wrong but Useful — What Covid-19 Epidemiologic Models Can and Cannot Tell Us. N Engl J Med 383, 303- 305 (2020).

18. Danillo Barros de Souza et al. Using discrete Ricci curvatures to infer COVID- 19 epidemic network fragility and

systemic risk. Journal of Statistical Mechanics: Theory and Experiment, Volume 2021, (2021).

19. Surkova, E., Nikolayevskyy, V. and Drobniewski, F. False-positive COVID-19 results: hidden problems and costs. The Lancet, Vol 8, Issue 12, 1167-1168, (2020).