Second Wave Analysis and Confirmed Forecasts of the SARS-Cov-2 Epidemic Outbreak in São Paulo, Brazil

Sergio Celaschi [1*]

[1] CTI Renato Archer, Rod. D Pedro I (SP-65), Km 143,6 - CEP 13069-901, Campinas – SP, 13069, Brazil.

Abstract

Objective: A SEIR compartmental model was previously selected to estimate future outcomes to the dynamics of the Covid-19 epidemic breakout in Brazil. *Method:* Compartments for individuals vaccinated and prevalent SARS-Cov-2 variants were not included. A time-dependent incidence weight on the reproductive basic number accounted for Non Pharmaceutical Interventions (NPI). A first series of published data from March 1st to May 8, 2020 was used to adjust all model parameters aiming to forecast one year of evolutionary outbreak. The cohort study was set as a city population-based analysis. *Analysis:* A population-based sample of 25,366 confirmed cases on exposed individuals was used during the first study period. The analysis was applied to predict the consequences of NPI enforcements followed by progressive releases, and indicates the appearance of a second wave starting last quarter of 2020. *Findings:* By March 1st 2021, the number of confirmed cases was predicted to reach 0.47Million (0.24-0.78), and fatalities would account for 21 thousand (12-33), 5 to 95% CRI. A second series of data published from May 9, 2020 to March 1st, 2021 confirms the forecasts previously reported for the evolution of infected people and fatalities. *Novelty:* By March 1st 2021, the number of confirmed cases reached 527,710 (12% above the predicted average of accumulated cases) and fatalities accounted for 18,769 (10% below the accumulated average of estimated fatalities). After March 1st, new peaks on reported numbers of daily new infected and new fatalities appeared as a combined result to the appearance of the prevalent SARS-CoV-2 P1 variant, and the increased number of vaccinated individuals.

Keywords: COVID-19; Brazil; Confirmed Forecast; NPI and Mitigation Policy; Second Wave; Prevalent Variants; Vaccination.

1. Introduction

COVID-19 was first reported in Brazil in February 2020 and one semester later, the country became one of the worst affected globally. After six months from the first reported case, the number of confirmed cases and deaths surpassed 3.9 million and 120 thousand, respectively; facts that question the availability of public health care for a major fraction of society [1, 2]. Since October 2020, all parts of the world were impacted by the COVID-19 epidemic, with more than 40 million cases and 1.1 million deaths reported [3]. This work aims to confirm previous forecasts reported on May, 2020 [4] for the COVID-19 epidemic expansion of in S. Paulo, the most populated Brazilian city with its 12.2 million inhabitants. This city was selected once the first positive tested patient in Brazil was there. Since then, by April 25 2021, were officially confirmed in Brazil 14.4 million cases with 391 thousand deaths and in S. Paulo city 704,989 cases and 26,642 deaths. Before the beginning of vaccination, the worldwide response to the pandemic has been the introduction of Non Pharmaceutical Interventions (NPIs) as mitigation policies.

Regarding predictions, epidemiological models are commonly stochastic, spatially diffusive, network based, with heterogeneous sub-populations (meta-population approaches) [5-9]. However, the compartmental models parameters are more directly related to and interpretable as physical processes [10-12]. On the other hand, deterministic models impose restrictive analysis, once the dynamics of the host population and the virus are not deterministic. The population has free will, and the virus undergoes "random" mutations [13, 14].

The primarily intent of this work was to build a simple epidemiological and compartmental model to predict and confirm the main results for the basic dynamics of the Covid-19 epidemic breakout in the target city. Compartments for vaccinated individual, and prevalent SARS-Cov-2 variants were not included in the model [15]. The infectious disease real-time transmissibility is often characterized by the instantaneous reproduction $R(t)$, the expected number of secondary infections caused by an infector within a short time window. Equivalently, $R(t)$ can be expressed as the transmission rate $\beta(t)$ divided by the rate γ_o into which infected people recover or die. Mitigation policies aim to control the outbreak reducing the $R(t)$ value. In this work, a time-varying incidence weight on the basic reproductive number R_o, was adopted to account for external influences as an elementary empirical approach. A first series of official published data from March 1st to May 8, 2020 [4] was used to adjust the model parameters, aiming to forecast one year of the COVID-19 evolutionary outbreak. The analysis was applied to predict the consequences of NPI enforcements followed by progressive releases and indicates the appearance of a second epidemic wave. A second series of official published data from May 9, 2020 to March 1st, 2021 confirms the forecasts previously reported for the evolution of infected people and fatalities associated to this epidemic outbreak in the city of S. Paulo.

2. Compartmental Model and Methodology

The compartmental SEIR model (Figure 1) represents one of the most adopted mathematical models to characterize an epidemic dynamics, and to predict possible contagion scenarios. This model can be useful to assess the effectiveness of various measures, such as social distancing, lock-down, mask wearing, closing public areas, among others. The model is based on a series of dynamic ordinary differential equations that consider the amount of the population subject to contagion, the trend over time of individuals who get exposed, recover after infection and the individuals who unfortunately die [12]. The coefficients of the equations represent the ratios of variation over time of different compartments, i.e., susceptible, exposed, infected, recovered, and dead. These coefficients have often been considered as constants. However, as constants they cannot take into account external influences, such as the NPIs, as social distancing, mask wearing, lockdowns, closed public places or the possible change in health conditions of infected individuals due to pharmacological development. In this approach, a time-dependent model parameter was introduced. It was assumed the infection rate time-dependent, considering that the number of contacts between people, during NPIs, changes proportionally to their overall mobility. Enforced NPIs in the city of S. Paulo reduced the growth of infected individuals after the third week (March 22, 2020). Additionally, the fatality rate μ was set constant, which is an oversimplification not supported by data. In Figure 1, the susceptible (S) is part of the population that could be potentially subjected to the infection. The exposed (E) is the fraction of the population that has been infected but is not infectious, what is called a latent phase. The infective (I) represents the infective hosts after the latent period. The recovered (R) accounts for the fraction of hosts after healing, R hosts are not reintroduced into the susceptible category assuming they became immune to the disease during the study period. The fatality (F) is the dead portion of the population N considered in this cohort study, $S(t) + E(t) + I(t) + R(t) + F(t) = N$. As can be noticed in this figure, compartments for vaccinated individuals and prevalent SARS-CoV-2 variants were not included in the model.
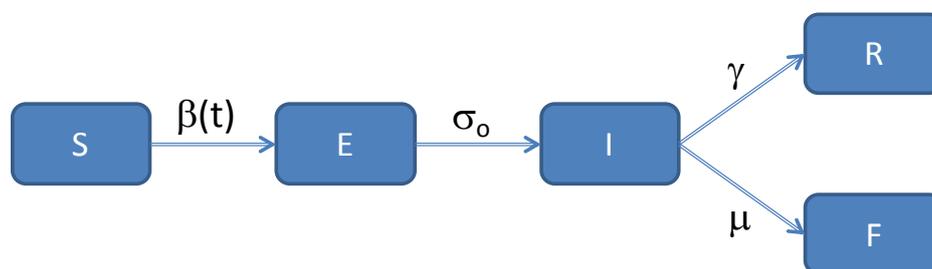
Figure 1. Compartmental epidemiological model used in this cohort study. The coefficients represent the ratio of variation of different compartments, i.e., susceptible, exposed, infected, recovered, and dead

The research methodology employed in this work is presented as a flowchart in Figure 2 for the largest populated city in Brazil (S. Paulo) selected in this cohort study. The basic concepts applied to the selected SEIR model are available elsewhere [4].
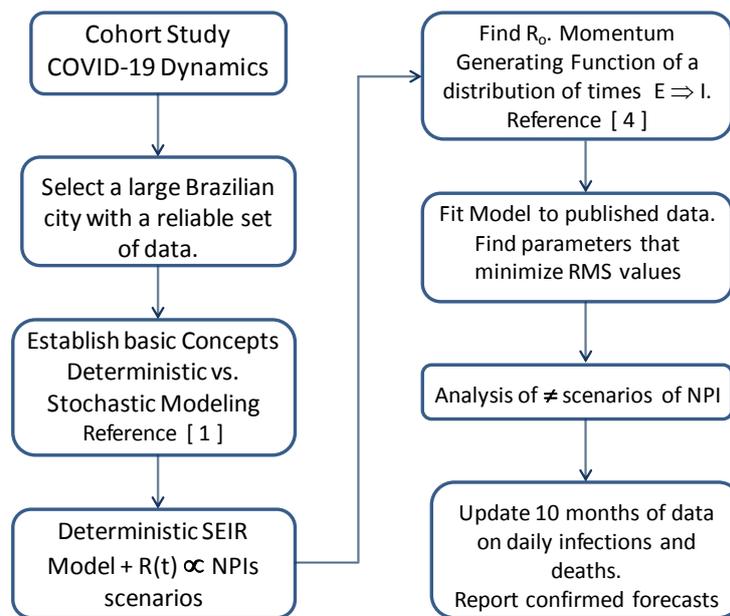
Figure 2. Flowchart for the research methodology applied to this work

A first data series published from March 1st to May 8, 2020 initially considered the epidemiological parameters that govern the selected compartmental model dynamics and adjust the model parameters aiming to forecast the COVID-19 evolutionary outbreak for one year. The model provides predictions of the time series of infected individuals and fatalities in the studied population. Simulations of average epidemic outbreak scenarios were done dependent on the level of NPIs enforcement and release. Forecasts pointed how such a policy alters the contamination pattern and suggested the existence of a second epidemic wave.

A key parameter in deterministic transmission models is the basic reproductive number R_o, which is quantified by both, the pathogen and the particular population in which it circulates. Thus, a single pathogen, like the SARS-CoV-2, will have different R_o values depending on the characteristics and transmission dynamics of the population experiencing the outbreak. The methodology to estimate $R_o = 2.53 \pm 0.05$ was presented in a previous publication [4]. Accordingly, $R(t)=\psi(t).R_o$ becomes dependent on the NPI policy $\psi(t)$. The analysis of a time-varying reproductive number $R(t)$ is applied to forecast the consequences of enforce, maintain and release the NPIs over specific time periods. By May 2020, during the first part of this study, without enough data on asymptomatic hosts, the percentage of symptomatic host's ξ_o was arbitrarily set as 50%. Later on, applying statistical analysis on published and reliable data, the percentage value of non-symptomatic and symptomatic hosts were estimated to be respectively (54±9) % and (46±9) % [14].
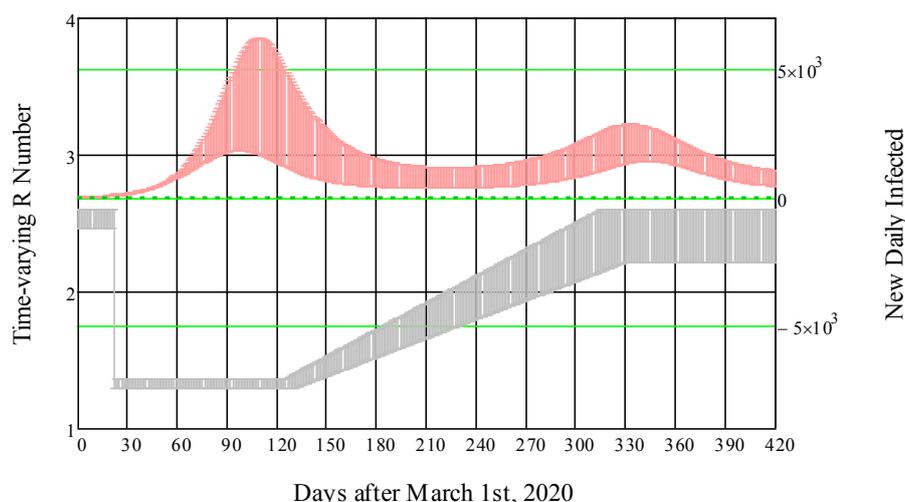
Figure 3. Long term prediction (red shaded area) of daily infected hosts (right vertical axis) as consequence of proposed variations of R(t) (left vertical axis - gray shaded area) modeling the NPI after March 22nd, 2020 in the city of S. Paulo. The values of the SEIR parameters are: Ro = 2.53 ± 0.05, $N_{average}$=8.5x10^5 (5x10^5 <N<14×10^5), β_o = 0.980±0.012, γ_o = 0.387±0.018, ψ_o = 0.525, μ = 0.015±0.001, σ_o = 0.5, ξ_o = 0.5.

Figure 3 illustrates the daily simulations of infected hosts as the result of the enforcement and the progressive release of NPIs during the period of study. The NPI ending time releases was selected to best fit the reported data on infected hosts. Surprisingly or not, the model reported on May, 2020 suggests the existence of a second Covid-19 outbreak wave, characterized by the presence of a new peak on the daily infected behavior, months after the first one. The number of infected individuals was estimated to be lower compared to the first outbreak. This is a consequence of the partial reduction in the initial number N of susceptible individuals.

3. Discussion and Forecasts

The fittings to the accumulated data on infected individuals and the fatalities selected by SEIR model are shown in Figure 4. A first series of official published data from March 1st to May 8, 2020 was previously used to adjust the model parameters, aiming to forecast about one year of the COVID-19 evolutionary outbreak (Figure 4a). The reported fitting values are: $5 \times 10^5 < N < 14 \times 10^5$, $R_o = 2.53 \pm 0.05$, $\beta_o = 0.980 \pm 0.012$, $\gamma_o = 0.387 \pm 0.018$, $\psi_o = 0.525$, $\mu = 0.015 \pm 0.001$, $\sigma_o = 0.5$. The cohort study was set as a city population-based analysis. The population-based sample of 25,366 individuals, during the initial study period corresponded to the number of confirmed cases on the exposed individuals. In the same period, the accumulated fatalities accounted for 2,110 confirmed deaths.
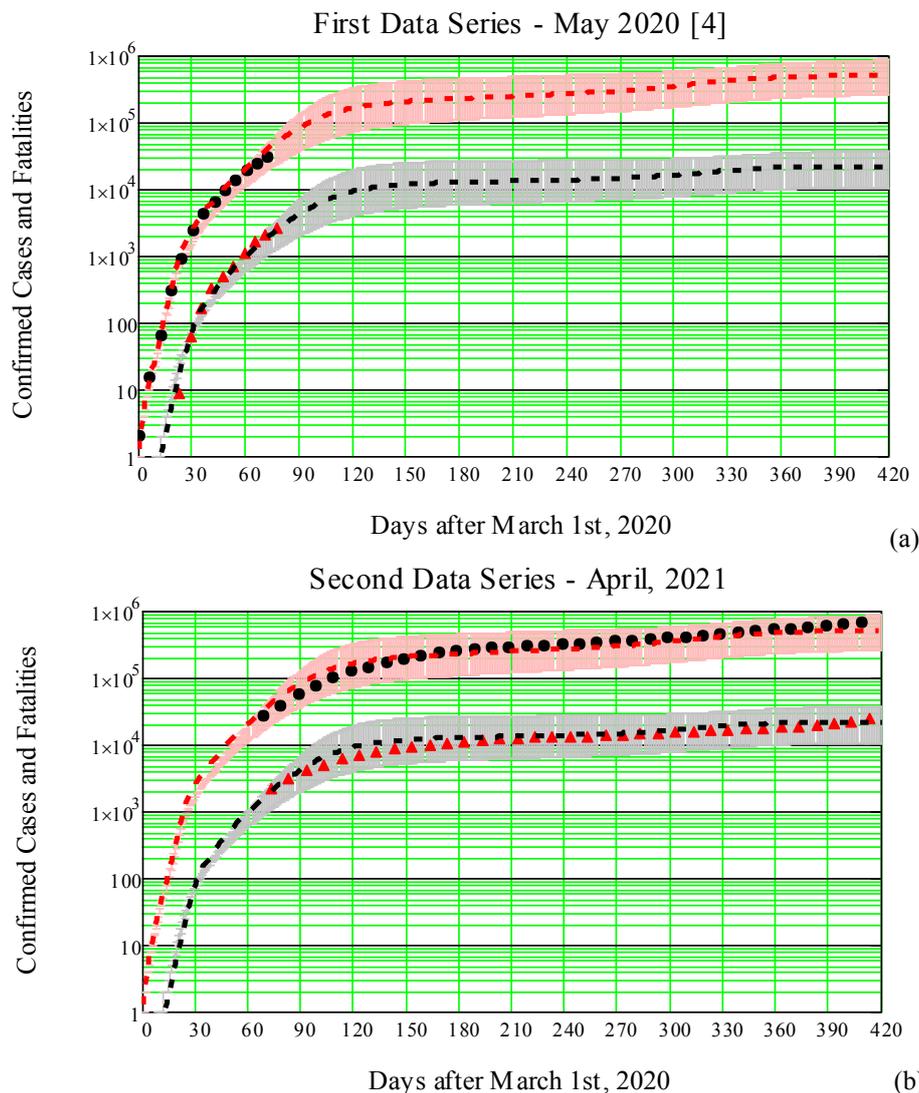
Figure 4. (a) Primary data on confirmed cases and fatalities are shown by solid black circles, and red triangles respectively. Shadow areas account for the 5-95% CRI on model parameters. The fitting mean values to model are: N = 8.5×10⁵, Rₒ = 2.53, βₒ = 0.980, γ = 0.387, μ = 0.015, ψₒ = 0.525, σₒ = 0.50, and t = 22 days. (b) Second series of official published data from May 9, 2020 to April 24, 2021 on confirmed cases and fatalities are shown by solid black circles and red triangles respectively.

A second series of official published data [16] from May 9, 2020 to April 24, 2021 shown in Figure 4b, confirms the forecasts reported previously on May 2020 for the infected people evolution and fatalities associated to this epidemic outbreak in the city of Sao Paulo. The average fitting value to confirmed cases (logarithmic scale) for the whole study period presents standard deviation SD = 0.07 and root mean square RMS = 0.08. By March 1st 2021, the official number of confirmed cases reached 509,193 a 20 fold increase on the population-based sample (25,366), with

12% deviation above the predicted average value. The reported fatalities accounted for 18,769 (10% below the average of predicted deaths). In short, the previous forecasts reported on May, 2020 for the expansion of the COVID-19 epidemic in the most populated Brazilian city (12.2 million inhabitants) are confirmed by the epidemiological SEIR model as shown in Figure 3b. During the period March 2nd to April 25, 2021, the last day reported, new sharp peaks appeared on the reported numbers of daily new infected and new fatalities. These are caused by a combined result to the appearance of the prevalent SARS-CoV-2 P1 variant and the increased number of vaccinated individuals, as commented later.
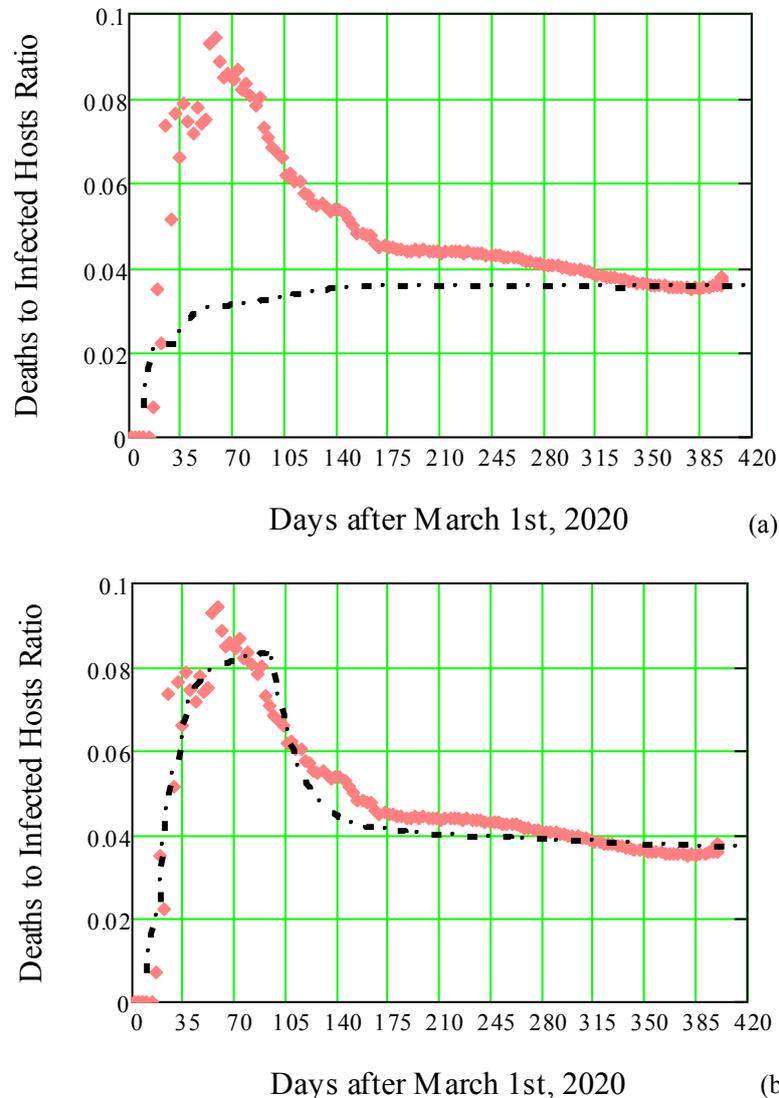
Figure 5. Ratio of deaths to confirmed cases (open red diamonds) as a function of time. (a) An assumed constant fatality rate in the compartmental model leads to a monotonic behavior in disagreement to data (shown by the black dash-and-dot line). (b) A time-varying fatality rate μ(t) (not included in the previous report) obtained by data fitting to the first 90 days after March 1st, 2020, would further improve the SD and RMS values.

The data ratio, fatalities to confirmed cases in S. Paulo, presents a temporal evolution that is not linear neither monotonic (Figure 5), as previously reported for the whole Brazilian country [14]. The non-zero ratio of deaths/ (confirmed cases) by COVID-19 (full red diamonds), begins two weeks after the first reported case, rises monotonically, after roughly seven weeks reaches a maximum of 0.085, and drops continuously to a ratio around 0.04. Possible explanations are many: Increased number of tests and reports available after the first weeks of the epidemic outbreak raises the number of confirmed cases; ICU improved medical procedures; among others. COVID-19 was detected in Brazil in the 9th EPI week of 2020 and testing procedures for the SARS-CoV-2 virus was effectively included in the surveillance four weeks later. Such a time dependent ratio was not included in the compartmental model. The fatality rate μ was set constant throughout the study period, which leads to a monotonic behavior as shown in Figure 4a by the black dash-and-dot line. This simplification leads, for the same time period, to an overestimation of the infected host as presented in Figure 4, and degrades the values of SD and RMS. Adding to model a time-varying fatality rate μ(t) by data fitting to the first 90 days after March 1st, 2020 (Figure 5b) would further improve the values of SD and RMS to 0.05 and 0.06 respectively.

Novel SARS-CoV-2 lineage P.1 first identified in Manaus has been associated with potentially higher transmission rates, and now it is widespread across all Brazilian regions [17, 18]. On the other hand, a preliminary study conducted by Institute Butantan [19], suggests the vaccination in Brazil by CoronaVac can neutralize variants P.1 and P.2 of SARS-CoV-2. P.1 was first worldwide reported on January 10, 2021 after been detected in four travelers from Brazil at Haneda airport in Tokyo. It is circulating widely in Brazil and is considered a "variant of concern" because it is more transmissible. Epidemiologists believe P.1 is one of the major causes for recent jump in COVID-19 cases and deaths, alongside the relaxing of mobility restrictions in and after the holiday season and the slow pace of vaccination. Since February 15, in a period of 12 weeks, the variant P.1 of Covid-19, which emerged in Manaus, increased its representativeness in the cases of the São Paulo city from 0% to 91%. The estimate was based on virus samples taken from patients at Hospital São Paulo and confirms the prevalent spread of the new lineage [20]. As the result of this explosive spread, during the period March 2nd to April 25, 2021, the last day reported in this study, sharp peaks appears on the reported numbers of daily new infected and new fatalities, as presented in Figure 5. These sharp peaks are caused by the combined results of the appearance of the prevalent SARS-CoV-2 P1 variant and the increased number of vaccinated individuals. As mentioned previously, compartments for individuals' vaccinated and prevalent SARS-Cov-2 variants were not included in this simple model and their combined effects were not predicted.

Regarding the reported values on daily new cases and fatalities, due to a large data scattering, a 7 days moving average was applied to smooth the data. Part of the scatter data is due to the unreported cases and deaths during weekends and holydays, which accumulates large values on coming working days. Assuming the variation of R(t) as already modeled for the NPI enforcement and posterior progressive release (Figure 3), the reported numbers of daily new cases and fatalities are presents in Figure 5 in comparison to model predictions. Reducing data scattering bring the data points closer to average values. The low agreement between data and model during the first outbreak wave is, in part, caused by the assumption of a constant fatality rate.
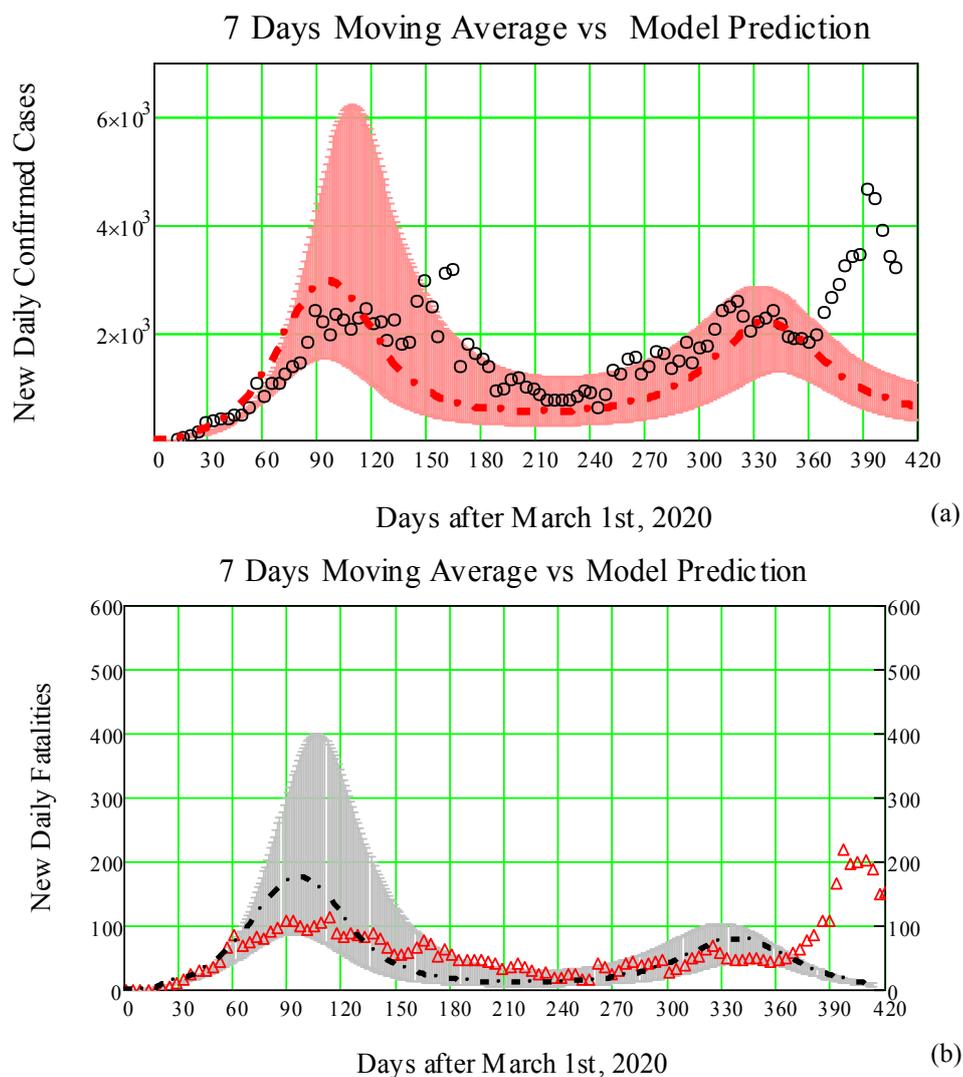
Figure 6. Reported (a) daily new cases (open black circle) and (b) fatalities (open red triangles) as a function of time compared to model predictions. The presence of a second wave becomes evident by published data and is confirmed by the SEIR model as shown in shaded areas.

This basic assumption leads to an overestimation of the first estimated peak of confirmed cases and fatalities. The presence of a second wave becomes evident by the data and model predictions presented by Figure 6. Surprisingly or not, the previous publication forecasted the presence of a second wave months before its arrival. As mentioned before, the predicted number of infected individuals within the second peak was estimated to be lower compared to the first one. The sharp peaks on confirmed cases and deaths at the end of the study period are the combined result of the appearance of the prevalent SARS-CoV-2 P1 variant and the increased number of vaccinated individuals.

4. Conclusion

This work confirms the forecasts previously reported on May, 2020 for the infected people evolution, fatalities and the appearance of a second wave starting in the last quarter of 2020 associated to the epidemic outbreak in S. Paulo, the largest Brazilian city [4]. The cohort study was set as a city population-based analysis in a sample of 25,366 confirmed COVID-19 cases on exposed symptomatic individuals. The analysis was applied to predict the consequences of the progressive NPI releases previously enforced during the first semester of 2020. An updated series of published data from May 9, 2020 to March 1st, 2021 confirms the forecasts previously reported for the infected people evolution and fatalities associated to this epidemic outbreak. By March 1st 2021, the official number of confirmed cases reached 527,710 a 20 fold increase on the population-based sample (25,366), with 12 % deviation above the predicted average value. The reported fatalities accounted for 18,769 (10% below the average of predicted deaths). During the period March 2nd to April 25, 2021, as a result of the explosive spread of the prevalent SARS-CoV-2 P1 variant and the increased number of vaccinated individuals, new sharp peaks appeared on the reported numbers of daily new infected and new fatalities. The simple SEIR model employed in this study did not include compartments for vaccinated individuals and prevalent SARS-Cov-2 variants. As a result, their combined effects were not predicted.

5. Declarations

5.1. Funding and Acknowledgements

5.2. Ethical Approval

The manuscript does not contain experiments on animals and humans; hence ethical permission was not required.

5.3. Data Availability Statement

The data presented in this study are available in Coronavirus casos em SP: https://www.seade.gov.br/coronavirus/.

5.4. Conflict of Interest

The author declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

6. References

[1] De Macedo, A. V. (2020). Brazil and COVID-19-A Fleeting Glimpse of What Is to Come. JAMA Health Forum, 1(9), e201061. doi:10.1001/jamahealthforum.2020.1061.

[2] De Souza, W. M., Buss, L. F., Candido, D. da S., Carrera, J.-P., Li, S., Zarebski, A. E., … Faria, N. R. (2020). Epidemiological and clinical characteristics of the COVID-19 epidemic in Brazil. Nature Human Behaviour, 4(8), 856–865. doi:10.1038/s41562-020-0928-4.

[3] Dong, E., Du, H., & Gardner, L. (2020). An interactive web-based dashboard to track COVID-19 in real time. The Lancet Infectious Diseases, 20(5), 533–534. doi:10.1016/s1473-3099(20)30120-1.

[4] Celaschi, S. (2020). Quantifying Effects, Forecasting Releases, and Herd Immunity of the Covid-19 Epidemic in S. Paulo – Brazil", The International Journal of Engineering and Science (IJES), 9(07), Series II, 33-42.

[5] Zlojutro, A., Rey, D., & Gardner, L. (2019). A decision-support framework to optimize border control for global outbreak mitigation. Scientific Reports, 9(1). doi:10.1038/s41598-019-38665-w.

[6] Wang, L., & Wu, J. T. (2018). Characterizing the dynamics underlying global spread of epidemics. Nature Communications, 9(1). doi:10.1038/s41467-017-02344-z.

[7] 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. doi:10.1126/science.abb3221.

[8] Parino, F., Zino, L., Porfiri, M., & Rizzo, A. (2021). Modelling and predicting the effect of social distancing and travel restrictions on COVID-19 spreading. Journal of the Royal Society Interface, 18(175). doi:10.1098/rsif.2020.0875.

[9] Eggo, R. M., Dawa, J., Kucharski, A. J., & Cucunuba, Z. M. (2021). The importance of local context in COVID-19 models. Nature Computational Science, 1(1), 6–8. doi:10.1038/s43588-020-00014-7.

[10] Kissler, S. M., Tedijanto, C., Goldstein, E., Grad, Y. H., & Lipsitch, M. (2020). Projecting the transmission dynamics of SARS-CoV-2 through the postpandemic period. Science, 368(6493), 860–868. doi:10.1126/science.abb5793.

[11] Godio, A., Pace, F., & Vergnano, A. (2020). SEIR Modeling of the Italian Epidemic of SARS-CoV-2 Using Computational Swarm Intelligence. International Journal of Environmental Research and Public Health, 17(10), 3535. doi:10.3390/ijerph17103535.

[12] Bastos, S. B., & Cajueiro, D. O. (2020). Modeling and forecasting the early evolution of the Covid-19 pandemic in Brazil. Scientific Reports, 10(1). doi:10.1038/s41598-020-76257-1.

[13] Korber, B., Fischer, W. M., Gnanakaran, S., Yoon, H., Theiler, J., Abfalterer, W., … Wyles, M. D. (2020). Tracking Changes in SARS-CoV-2 Spike: Evidence that D614G Increases Infectivity of the COVID-19 Virus. Cell, 182(4), 812–827.e19. doi:10.1016/j.cell.2020.06.043.

[14] Celaschi, S., "The Impact of SARS-CoV-2 Variant to COVID-19 Epidemic in Brazil", medRxiv. doi:10.1101/2020.09.25.20201558.

[15] Oliveira, J. F., Jorge, D. C. P., Veiga, R. V., Rodrigues, M. S., Torquato, M. F., da Silva, N. B., … Andrade, R. F. S. (2021). Mathematical modeling of COVID-19 in 14.8 million individuals in Bahia, Brazil. Nature Communications, 12(1). doi:10.1038/s41467-020-19798-3.

[16] Coronavirus casos em SP. Available online: https://www.seade.gov.br/coronavirus/ (accessed on April 25, 2020).

[17] Lamarca, A. P., de Almeida, L. G. P., Francisco, R. da S., Lima, L. F. A., Scortecci, K. C., Perez, V. P., … Vasconcelos, A. T. R. (2021). Genomic surveillance of SARS-CoV-2 tracks early interstate transmission of P.1 lineage and diversification within P.2 clade in Brazil. medRxiv. doi:10.1101/2021.03.21.21253418.

[18] Faria, N. R., Mellan, T. A., Whittaker, C., Claro, I. M., Candido, D. da S., Mishra, S., … McCrone, J. T. (2021). Genomics and epidemiology of a novel SARS-CoV-2 lineage in Manaus, Brazil. medRxiv. doi:10.1101/2021.02.26.21252554

[19] Palacios, R., Batista, A. P., Albuquerque, C. S. N., Patiño, E. G., Santos, J. D. P., Tilli Reis Pessoa Conde, M., ... & Kallas, E. G. (2021). Efficacy and safety of a COVID-19 inactivated vaccine in healthcare professionals in Brazil: the PROFISCOV study. SSRN Preprint. doi:10.2139/ssrn.3822780.

[20] Study suggests the Brazilian variant emerged in November, is more transmissible and can cause reinfection. Available online: https://agencia.fapesp.br/study-suggests-the-brazilian-variant-emerged-in-november-is-more-transmissible-and-can-cause-reinfection/35414/ (accessed on April 25, 2020).