Confirmed forecasts for the expansion of the COVID-19 epidemic in the S. Paulo city, Brazil

Sergio Celaschi*

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

**Correspondent author: sergio.celaschi@cti.gov.br*

Abstract

An epidemiological compartmental model reported on May, 2020 was previously selected to estimate future outcomes to the dynamics of the Covid-19 epidemic breakout in the city of S. Paulo, Brazil. A time-dependent incidence weight on the reproductive basic number accounted for the Non Pharmaceutical Interventions (NPIs). A first series of official published data from March 1st to May 31st, 2020 was used to adjust the model parameters aiming to forecast one year of the COVID-19 evolutionary outbreak. The cohort study was set as a city population-based analysis. The population-based sample, 25,366 during the study period, was the number of confirmed cases on exposed individuals. The analysis was applied to predict the consequences of releasing the NPIs, and indicated the appearance of a second wave starting last quarter of 2020. By January 31st 2021, the number of confirmed cases was predicted to reach 0.49 Million (0.28-0.77), and fatalities would account for 22 thousand (11-32), 5 to 95% CRI. A second series of official data published from June 1st 2020 to January 31st, 2021 confirms all forecasts previously reported for the evolution of infected people and fatalities associated to this epidemic outbreak in the city of S. Paulo. By January 31st 2021, the official number of confirmed cases reached 469,657 (4 % above predicted average of accumulated cases), and fatalities accounted for 17,333 (19% above accumulated average of fatalities).

Keywords: COVID-19; S. Paulo; confirmed forecast; NPI and mitigation policy; second wave.

1. Introduction

This work aims to confirm previous forecasts reported on May, 2020 [1] for the expansion of the COVID-19 epidemic in S. Paulo, the most populated Brazilian city with its 12.2 million inhabitants. It confirms all forecasts previously reported for the evolution of infected people and fatalities associated to this epidemic outbreak in town. The city of S. Paulo was selected once the first patient in Brazil was tested positive there. Since then, by January 12, 2021, were officially confirmed in Brazil 8.2 million cases with 205 thousand deaths, and in S. Paulo city 424,252 cases, and 16,208 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, diffusive-spatial, network based, with heterogeneous sub-populations (meta-population approaches) [2, 3, 4]. However, the parameters of compartmental models are more directly related to and interpretable as physical processes [5,6]. 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 [7,8].

The primarily intent of this work was to build a simple epidemiological tool to predict and confirm the main results for the basic dynamics of the Covid-19 epidemic breakout in the city of S. Paulo, Brazil. A time-varying incidence weight on the basic reproductive number R_o was used to take into account external influences, and quantify local NPIs. A first series of official published data from March 1st to May 31st, 2020 [9] 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 progressive releases,

and indicates the appearance of a second wave. A second series of official published data from June 1st 2020 to January 31st, 2021 [10] confirms all 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 (Fig. 1) represents one of the most adopted mathematical models to characterize an epidemic dynamics, and to predict possible contagion scenarios. The SEIR 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. 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 is 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 after the third week (March 22, 2020). Additionally, the fatality rate μ was set constant, which is an oversimplification not supported by data.

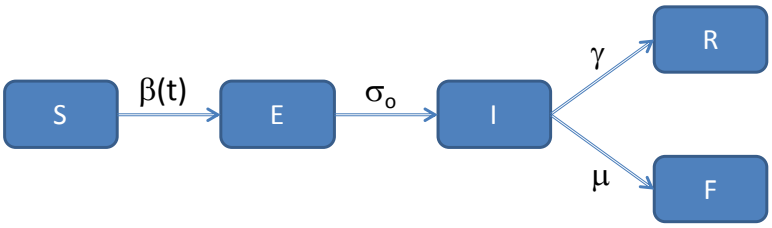


Figure 1. The susceptible (S) is part of the population that could be potentially subjected to the infection: at the initial time. The exposed (E) is the fraction of the population that has been infected but is not infectious: it can be called a latent phase. The infective (I) represents the infective hosts after the latent period. The recovered (R) accounts for the fraction of hosts after healing, and they are generally not reintroduced into the susceptible category if it is supposed that they became immune to the disease. The fatality (F) is the dead part of the population.

$$S(t)+E(t)+I(t)+R(t)+F(t)=N$$

The compartmental model and methodology used in this work was presented with additional details in a previously publication [1]. A first series of available published data from March 1st to May 31st, 2020 was initially considered to estimate the epidemiological parameters that govern the selected compartmental model dynamics, and adjust the model parameters aiming to forecast one year of the COVID-19 evolutionary outbreak. The model provides predictions of the time series of infected individuals and fatalities in the studied area. Simulations of midi-term scenarios of the epidemic outbreak

were done dependent on the level of NPIs enforcement and release. Forecasts showed how such a policy alters the pattern of contamination, 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. Accordingly, $R(t)=\psi(t).R_o$ becomes dependent on the NPI policy $\psi(t)$. The effectiveness of those NPIs may be quantified by a social distancing factor defined as $SD = 1 - \psi(t)$, as previously reported [1]. The analysis of a time-varying reproductive number $R(t)$ is applied to forecast the consequences of releasing the NPIs over specific periods of time.

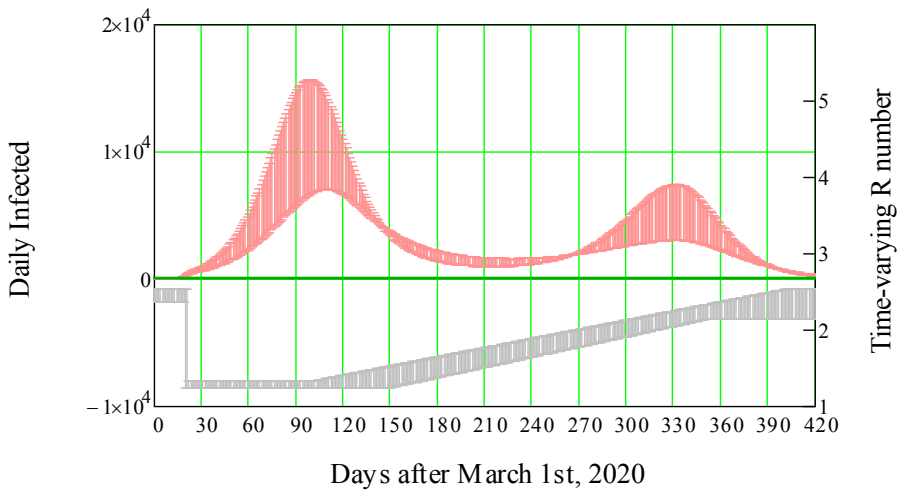
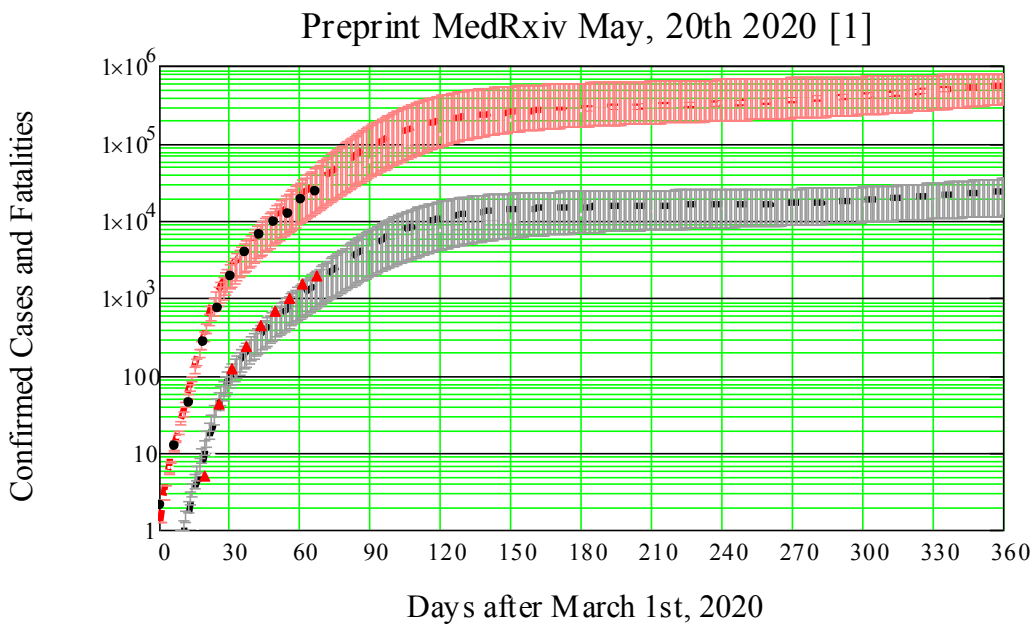


Figure 2. (a) Long term prediction (red shaded area) of daily infected hosts as previously reported as a consequence of proposed variations of R(t) (gray shaded area) modeling the NPI after March 22nd, 2020 in the city of S. Paulo.

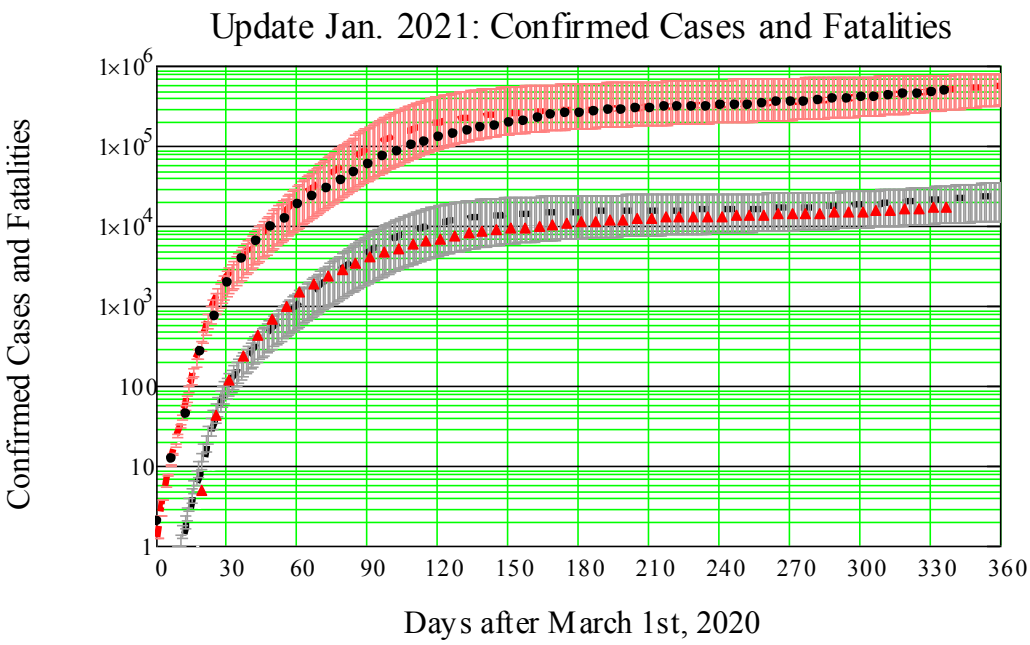
Figure 2 illustrates a simulation of daily infected hosts as the result of the enforcement and progressive realize of NPIs during the period of study. Surprisingly or not, the model reported on May, 2020 suggests the existence of a second outbreak wave of Covid-19 characterized by the presence of a new peak of daily infected appearing 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 data on infected individuals, and fatalities by the selected model are shown in Fig. 3a. A first series of official published data from March 1st to May 31st, 2020 was previously used to adjust the model parameters aiming to forecast about one year of the COVID-19 evolutionary outbreak. The reported fitting values are: $N = (5 \text{ to } 15).10^5$, $R_o = 2.53\pm0.05$, $\beta_o = 0.960\pm0.015$, $\gamma_o = 0.379\pm0.005$, $\psi_o = 0.525$, $\mu = 0.014$, $\sigma_o = 0.5$.



(a)

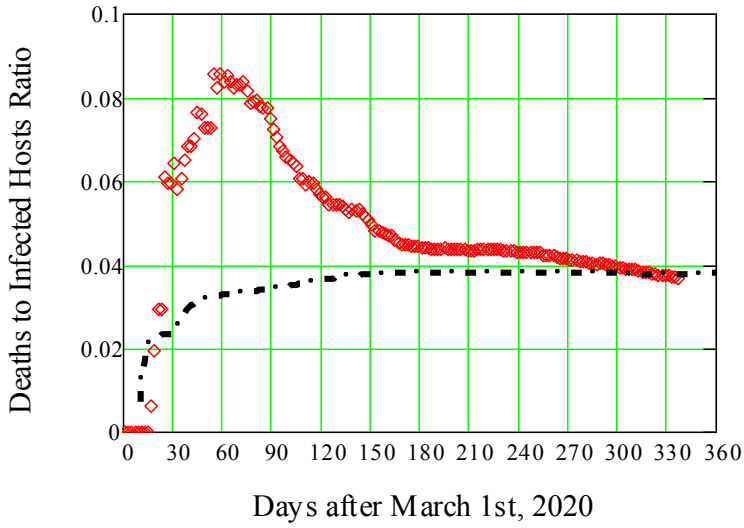


(b)

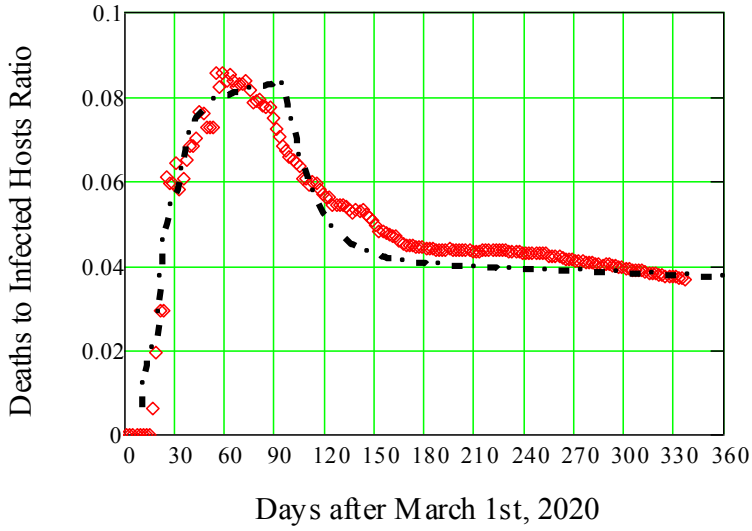
Figure 3. (a) Primary official data on confirmed cases and fatalities are shown by solid black circles, and red triangles respectively. Error bars account for the RMS values on model´s parameters. Average values and error bars for infected individuals, and fatalities are shown in red, and black dashed vertical lines respectively. The fitting mean values to model are: $N = 9.25x10^5$, $R_o = 2.53$, $\beta_o = 0.960$, $\gamma = 0.379$, $\mu = 0.017$, $\psi_o = 0.525$, $\sigma_o = 0.50$, and $t = 22$ days. (b) Second series of official published data from June 1st 2020 to January 12, 2021 [10] confirms all forecasts reported previously on May 2020 for the evolution of infected people and fatalities associated to this epidemic outbreak in the city of Sao Paulo, Brazil. The fittings for last 50 days of cases and fatalities are shown in the inset as an expanded view.

A second series of official published data from June 1st 2020 to January 12, 2021 [10] shown in Fig. 3b, confirms all forecasts reported previously on May 2020 for the evolution of infected people and

fatalities associated to this epidemic outbreak in the city of Sao Paulo, Brazil. 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 = 0.09. By December 31st 2020, the official number of confirmed cases reached 401,718, an increase over 18 folds on the population-based sample (25,366), with only 4 % variation above the predicted average value. The reported fatalities accounted for 15,679 (19% above the average of predicted deaths). In short, all previous forecasts reported on May, 2020 for the expansion of the COVID-19 epidemic in the most populated Brazilian city with its 12.2 million inhabitants are confirmed as shown in Fig. 3b.



(a)



(b)

Figure 4. 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 is shown by the dark brown dash-and-dot line. (b) A time-varying fatality rate $\mu(t)$ (not included in the previous report) obtained by data fitting to the first 90 days after March 1st, 2020, would further improve the values SD and RMS.

As reported elsewhere for the whole Brazilian country [8], the data ratio, fatalities to confirmed cases in S. Paulo, present a temporal evolution that is not linear neither monotonic (Fig. 4). The non-zero ratio of deaths/(confirmed cases) by COVID-19 (open red circles), 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 Fig.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 Fig. 3, and degrades the values of fitting parameters 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 (Fig. 4b) would further improve the values of SD and RMS to 0.05 and 0.06 respectively.

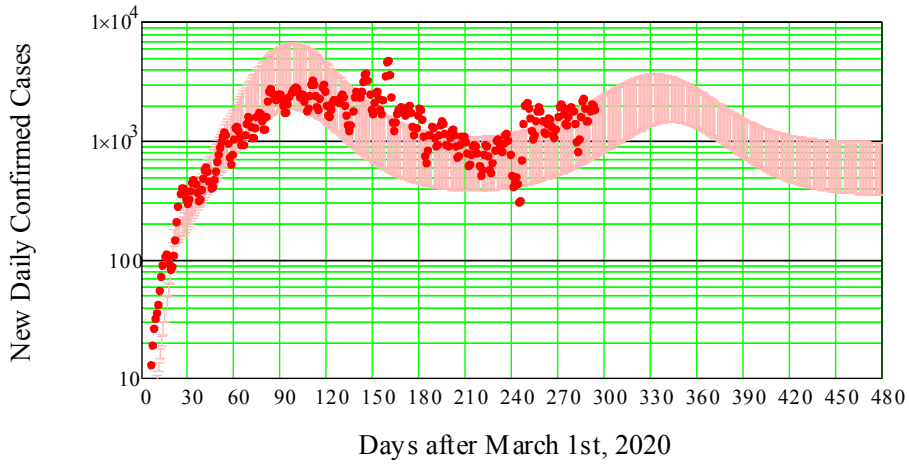
Regarding data on daily new cases and fatalities, due to a large data scattering, a smoothing data formula using a Gaussian Kernel Method - Equations (1) and (2), available in the Math Software for Engineering Calculations [11], within a four day (b = 4) period was applied to data on new cases and fatalities.

$$vy_i = \frac{\sum_{j=1}^{n} K\left(\frac{vx_i - vx_j}{b}\right).vy_j}{\sum_{j=1}^{n} K\left(\frac{vx_i - vx_j}{b}\right)} \qquad (1)$$

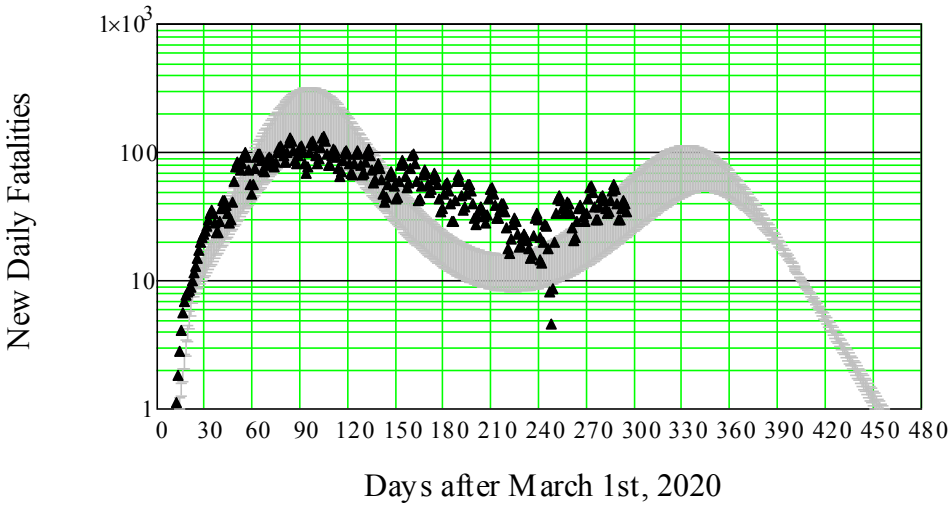
where

$$K(t) = \frac{1}{\sqrt{2\pi}.(0.37)}.e^{-0.5*\left(\frac{t}{0.37}\right)^2} \qquad (2)$$

Assuming the variation of R(t) as already modeled for the NPI progressive release, a comparison to the reported numbers of daily new cases and fatalities is presents in Fig. 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.



(a)



(b)

Figure 5. Reported (a) daily new cases (solid red circles) and (b) fatalities (solid black triangles) as a function of time compared to model predictions. The presence of a second wave becomes evident by published data and model predictions shown as 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 Fig. 5a and 5b. Surprisingly or not, the previous publication reported on the presence of a second peak months after the first one. As mentioned before, the predicted number of infected individuals within the second peak was estimated to be lower compared to the first one.

4. Conclusions

This work confirms all forecasts previously reported on May, 2020 for the evolution of infected people and fatalities associated to this epidemic outbreak in the city of S. Paulo, Brazil. The cohort study was set as a city population-based analysis in a sample of 25,366 confirmed COVID-19 cases on exposed individuals. The analysis was applied to predict the consequences of progressive releases of the NPIs enforced during the first semester of 2020. It forecasted the appearance of a second wave starting last quarter of 2020 to fadeout by May 2021. By January 31st 2021, the number of confirmed cases was predicted to reach 0.49 Million, and fatalities would account for 17 thousand. A second series of official data published later from June 1st 2020 to January 31st, 2021 confirms all forecasts previously reported for the evolution of infected people and fatalities associated to this epidemic outbreak in town. By January 31st 2021, the official number of confirmed cases reached 0.47 Million – 4 % above predicted average of accumulated cases), and fatalities accounted for 21 thousand – 19 % above the accumulated average of fatalities. In conclusion, all previous forecasts reported on May, 2020 for the expansion of the COVID-19 epidemic in S. Paulo, the most populated Brazilian city where the pandemic started in the country, are confirmed here.

5. Acknowledgements

The author acknowledges the financial grant by FUNTTEL. Financial grant # 01.16.0053.01 FINEP/MCTI under the Brazilian Ministry of Science, Communications, and Innovation. This paper reflects only the author's views and the Agencies are not responsible for any use that may be made of the information contained therein.

6. References

[1] Celaschi, S., *"Quantifying Effects, Forecasting Releases, and Herd Immunity of the Covid-19 Epidemic in S. Paulo – Brazil"*, The International Journal of Engineering and Science (IJES), Volume 9, Issue 07, Series II, pp 33-42, 2020. doi: 10.9790/1813-0907023342.

[2] Zlojutro, A., Rey, D. & Gardner, L., *"A decision-support framework to optimize border control for global outbreak mitigation"*, Nature Sci Rep 9, 2216 (2019). doi.org/10.1038/s41598-019-38665-w.

[3] Wang, L., Wu, J.T., *"Characterizing the dynamics underlying global spread of epidemics".* Nature Commun 9, 218 (2018). doi: 10.1038/s41467-017-02344-z

[4] Ruiyun Li, Sen Pei et. Al*., "Substantial undocumented infection facilitates the rapid dissemination of novel coronavirus (SARS-CoV-2)*", Science 01: Vol. 368, Issue 6490, pp. 489-493, May 2020. doi: 10.1126/science.abb3221

[5] Kissler, S. M., Tedijanto, C., Goldstein, E., ET. Al., *"Projecting the transmission dynamics of SARS-CoV-2 through the postpandemic period"*, Science 14 Apr 2020. doi: 10.1126/science.abb5793

[6] Godio A, Pace F, Vergnano A. SEIR Modeling of the Italian Epidemic of SARS-CoV-2 Using Computational Swarm Intelligence. Int J Environ Res Public Health. 2020;17(10):3535. Published 2020 May 18. doi:10.3390/ijerph17103535

[7] B. Korber, B. et al*., "Tracking changes in SARS-CoV-2 Spike: evidence that D614G increases infectivity of the COVID-19 virus",* Journal Pre-proof, Reference: CELL 11502, https://doi.org/10.1016/j.cell.2020.06.043

[8] Celaschi, S., *"The Impact of SARS-CoV-2 Variant to COVID-19 Epidemic in Brazil"*, Preprint MedRxiv. doi: https://doi.org/10.1101/2020.09.25.20201558

[9] Coronavirus casos em SP https://www.seade.gov.br/coronavirus/. Page visited on May, 2020.

[10] http://www.saopaulo.sp.gov.br/wp-content/uploads/2021/01/20210131_dados_covid_municipios_sp.csv. Page visited on Jan. 31, 2021.

[11] https://www.mathcad.com/en/, Page visited on Jan. 12, 2021.