

Stochastic analysis using public data for forecasting of epidemic spreading of the novel coronavirus disease

Leonardo dos Santos Lima (✉ lslima7@cefetmg.br)

CEFET-MG <https://orcid.org/0000-0003-0410-3640>

Research Article

Keywords: Stochastic, coronavirus, spreading

Posted Date: September 17th, 2020

DOI: <https://doi.org/10.21203/rs.3.rs-78321/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Stochastic analysis using public data for forecasting of epidemic spreading of the novel coronavirus disease

Leonardo S. Lima¹

¹Federal Education Center for Technological of Minas Gerais, 30510-000, Belo Horizonte, MG, Brazil.

*lslima@cefetmg.br

ABSTRACT

We propose a stochastic model for epidemic spreading of the novel coronavirus based in data supported by the Brazilian health agencies. Furthermore, we performed an analysis using the Fokker-Planck equation estimating the novel cases in the day t as the mean half-width of the distribution of novel cases $P(N, t)$. Our results display that the model based in the Itô diffusion adjusts well to the results supplied by health Brazilian agencies due to large uncertain in the official data and to the low number of tests realized in the population.

Introduction

The use of models and mathematical methods for theoretical researchers to the study the spread of contagious diseases goes back a least to some works by Daniel Bernoulli in XVIII century on smallpox.¹ One of these powerful techniques are nonlinear differential equations (NL) which are an important topic in different branches of the science and engineering due to their ability to explain several complex behaviors in the nature.¹⁻⁵

Diseases as the typhoid fever, smallpox and also the COVID-19 are spread largely by individuals who can transmit the disease but who do not exhibit overt symptoms. Considering x and y the proportions of susceptible and carriers, respectively, in the population, suppose that carriers are identified and removed from the population at a rate $dy/dt = -\alpha y$. If the disease spreads at a rate proportional to the product x and y , then we will have $dx/dt = -\gamma xy$. We can solve the set of equations for y at instant t subject to initial condition $y(0) = y_0$ to obtain the proportion of the population that escapes of the pandemic as the the limit value of x as $t \rightarrow \infty$ from the initial condition $x(0) = x_0$. For the smallpox, once contracted and survived, the individual present a permanent immunity. Thus considering a group of individuals born in a given time t and considering $n(t)$ the number of the individuals surviving a time t later. We make $x(t)$ the number of members of this group who do not have smallpox at time t and who are therefore susceptible and we let α be the rate at which susceptibles contract smallpox. Furthermore, we let v , the rate at which people who contract smallpox die from the disease. Finally, We let $\mu(t)$ be the death rate from all cases other than smallpox. Then dx/dt (the rate at which the number of susceptible declines), is given by $dx/dt = -[\alpha + \mu(t)]x$, where the first term is the rate at which the susceptibles contract smallpox, and the second term is the rate at which they die from all other cases. Furthermore $dn/dt = -v\alpha x - \mu(t)n$, where dn/dt is the death rate. The two terms on the right side are the death rates due to all other causes, respectively. In the end, we define the variable z given by $z = x/n$ and obtain that z satisfies the following nonlinear differential equation $dz/dt = -\alpha z(1 - vz)$, with the initial condition $z(0) = 1$ which is independent of $\mu(t)$. The v and α parameters were estimated as $v = \alpha = 1/8$. From these values, one can determine the proportion of individuals who never had acquired smallpox.

We can obtain a more realistic model adding some randomness in the differential equation above, where the size of the population at time t and $\alpha(t)$, the relative rate of growth as a function of time t , $\alpha(t)$ are not deterministic, but subject to some random due to environmental noise, so that, $\alpha(t) = r(t) + \text{"noise"}$. The exact behaviour of the 'noise term' is not known. However, its probability distribution, yes. The function $r(t)$ is assumed to be non random.⁶⁻⁸

In this work, we investigate the Itô diffusion model with additive white noise and nonlinear terms as a possible model for the spread of the COVID-19 in Brazil. Due to uncertainly in the official data about the real cases numbers generated by the low number of tests made in the population generates a large randomness in the data and therefore, makes the use of the stochastic analysis greatly adequate to treat the spread of time evolution of the COVID-19. Furthermore, we derive the stochastic differential equation, in Itô calculus, corresponding to nonlinear Fokker-Planck equation derived in framework of the non-additive statistical mechanics, which must also obey to stylized features of the financial market as inverse cubic law⁹⁻¹⁷.

The plan of this paper is the following. In section 1, we describe the stochastic model. In section 2, we make the connection

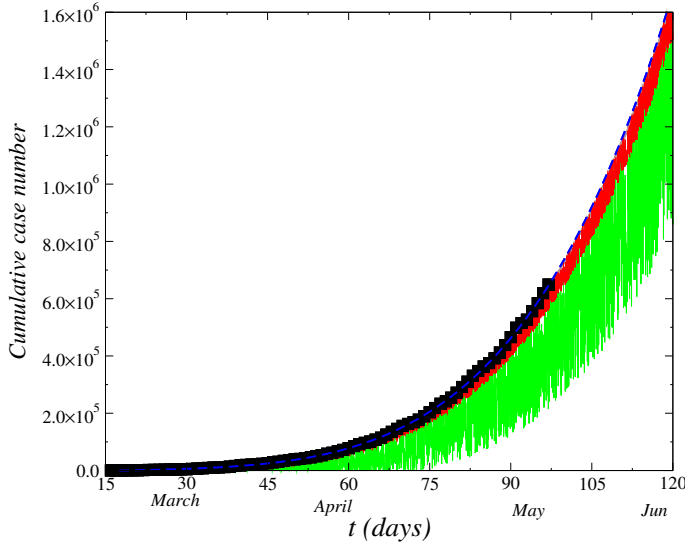


Figure 1. Effect of the stochastic term on cumulative total case number of the COVID-19 in Brazil. We perform the calculations for different values of β_0 in model Eq. (1) as $\beta_0 = 0.0$ (dashed blue-line), $\beta_0 = 3.0 \times 10^{-6}$ (red-line) and $\beta_0 = 5.0 \times 10^{-6}$ (green-line).

between stochastic differential equation and non linear Fokker-Planck equation. In section 3, we present the analytical results by Fokker-Planck equation. In section 4, we present our conclusions and final remarks.

1 A phenomenological Langevin equation for dynamics of the COVID-19

The behavior of the cumulative total cases number $\mathcal{P}(t)$ of infected by coronavirus registered in the Brazil as function of time (days) from 15th March, 2020 is displayed in Fig. 2. From fit of least minimums squares to the set of officials data of the Brazilian ministry of health, we estimate the curve for the behavior of the total cases number of COVID-19 given by a polynomial of fourth degree of t (dashed-blue-line) in the Fig. 2 given as $f(t) = -4397.7 - 34.3t + 26.2t^2 - 0.8t^3 + 0.01t^4$. The data were registered in the period 03 - 15 - 06 - 05. The points of non differentiability is due to an uncertainty of the official data and to population isolation conditions. Hence, for modeling of this behavior, we add a random term together with nonlinear terms in Eq. (1) with aim to simulate the effect of this uncertainty.

$$d\mathcal{P}(t) = (f(t) + A(\mathcal{P}(t), t))dt + B(\mathcal{P}(t), t) \circ dW \quad (1)$$

where $f(t)$ is a polynomial of least squares fit ($n = 4$ degree). In Fig. 1, we obtain the behavior of $f(t)$ given by $f(t) = -4397.7 - 34.3t + 26.2t^2 - 0.8t^3 + 0.01t^4$. Furthermore, we have $A(\mathcal{P}(t), t)$, given by Bernoulli's model: $A(\mathcal{P}(t), t) = \alpha\mathcal{P}(t)(1 - \nu\mathcal{P}(t))$ and $B(\mathcal{P}(t), t) = \beta_0 t^3$ (multiplicative white noise) in Eq. (1) and we perform the calculation for different values of β_0 (constant). We obtain a strong oscillation of the curve with the increasing of β_0 indicating so, an increase of growing of the uncertainty of the cumulative total case number. The data range considered is from 15-March until 06-June.

2 Nonlinear Fokker-Planck equation

2.1 Nonlinear Fokker-Planck equation within a Itô prescription

The nonlinear Fokker-Planck equation is given in the non-additive statistical mechanics by¹⁸

$$\frac{\partial P(x, t)}{\partial t} = D \frac{\partial^2 \{ [P(x, t)]^{2-q} \}}{\partial x^2} - \frac{\partial [P(x, t)K(x, t)]}{\partial x}. \quad (2)$$

We obtain that Eq. (2) is equivalent to following Itô stochastic differential equation

$$dX = A(X(t), t)dt + \phi(X(t), t) \circ dW, \quad (3)$$

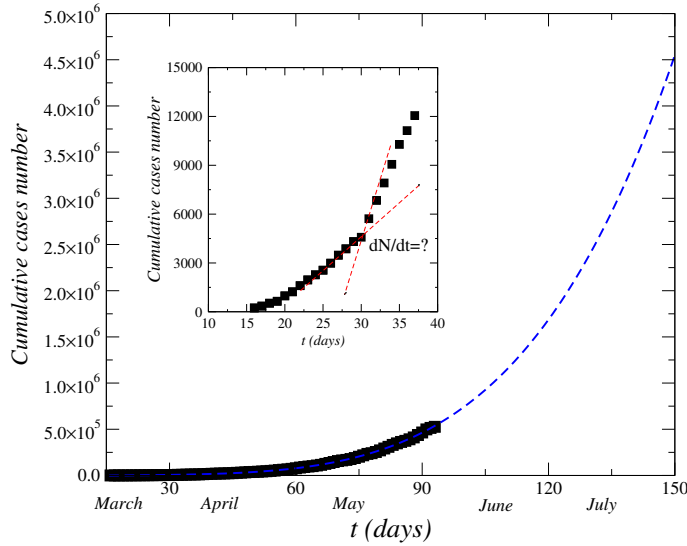


Figure 2. Total novel cases number fit of COVID-19 $\mathcal{P}(t)$. The squares (in black) are the cumulative case number registered in the Brazil in the range considered. We obtain a polynomial adjustment in the range considered (dashed-blue-line). In the inset, we display the non differentiability of the data curve.

where $A(X(t), t) = K(X(t), t)$ and $\phi(X(t), t) = [P(X(t), t)]^{\frac{1-q}{2}}$. $X(t)$ is a stochastic process defined on probability space Ω , being the triple (Ω, \mathcal{F}, P) a probability space, where \mathcal{F} is a σ -algebra and P a probability measure that is, a function that to every set $A \in \mathcal{F}$ assigns a number in range $[0, 1]$, where $P(\Omega) = 1$, $P(\emptyset) = 0$ and

$$P\left(\bigcup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} P(A_n). \quad (4)$$

The random variable X , defined on Ω with the property that for every Borel subset B of \mathbb{R} , the subset of Ω given by $\{X \in B\} = \{\omega \in \Omega; X(\omega) \in B\}$ is in the σ -algebra \mathcal{F} . Moreover, $P(X(t), t)$, given X a random variable on a probability space (Ω, \mathcal{F}, P) is the probability measure of X . μ_X assigns to each Borel subset B of \mathbb{R} the mass $\mu_X(B) = P\{X \in B\}$ ^{7,8}. Furthermore, we have that dW is the Wiener increment, where $W(t)$ is a Wiener process also known as Brownian motion. We have an environment stochastic white noise $\zeta(t)$ that if relate with the Wiener process $W(t)$ by $W(t) = \int_0^t \zeta(t') dt'$. Even though in the literature had been explained differently¹⁹, we have that $W(t)$ which is the integral of $\zeta(t)$, is not the derivative of the Wiener process, $\zeta(t) \neq dW(t)/dt$, being therefore $W(t)$ not differentiable⁶

$$\frac{dW(t)}{dt} = \lim_{\Delta t \rightarrow 0} \frac{\Delta W(\Delta t)}{\Delta t} \sim \frac{\sqrt{\Delta t}}{\Delta t} \rightarrow \infty. \quad (5)$$

The Itô integral for ϕ is given by

$$\int \phi(X(t), t) dW(t) = \text{ms-lim} \left\{ \sum_{j=1}^n \phi(t_{j-1}) [W(t_j) - W(t_{j-1})] \right\}, \quad (6)$$

where ms-lim means square limit. $W(t)$ is a Markovian process, presenting normal distribution, satisfying the conditions $\langle \zeta(t) \rangle = 0$ and $\langle \zeta(t) \zeta(t') \rangle = \delta(t - t')$.

We can write the Eq. (2) as

$$\frac{\partial P(x, t)}{\partial t} = D \frac{\partial^2 \{ [P(x, t) P(x, t)]^{1-q} \}}{\partial x^2} - \frac{\partial [P(x, t) K(x, t)]}{\partial x}, \quad (7)$$

where we define $A(x, t) = K(x, t)$ and $\phi(X(t), t) = [P(X(t), t)]^{\frac{1-q}{2}}$ to obtain the correspondent Itô stochastic differential equation

$$dX = K(X(t), t) dt + [P(X(t), t)]^{\frac{1-q}{2}} \circ dW, \quad (8)$$

which is the Eq. (3). The solution for $X(t)$ using the Stratonovich integral is given as

$$X(t) = X(0) + \int_0^t K(s, X_s) ds + \int_0^t \phi(s, X_s) dW_s. \quad (9)$$

This implies that $X(t)$ is the solution of the following modified Itô equation

$$X(t) = X(0) + \int_0^t K(s, X_s) ds \quad (10)$$

where ϕ' denotes the derivative of $\phi(x, t)$ with respect to x . Therefore, the Eq. (2) in Itô calculus is different of the Stratonovich interpretation.

From the Feynman-Kac theorem, let $h(x)$ be a Borel-measurable function. Fix $T > 0$, and let $t \in [0, T]$ be given, we define the function

$$g(x, t) = E^{t,x} h(X(T)), \quad (11)$$

where

$$E[g(X)] = \int_{-\infty}^{\infty} |g(x)| P(x) dx. \quad (12)$$

Furthermore, we assume $E^{t,x} |h(X(T))| < \infty$ for all t and x , then $g(x, t)$ satisfies the partial differential equation

$$\frac{\partial g(x, t)}{\partial t} + K(x, t) \frac{\partial g(x, t)}{\partial x} + \frac{1}{2} \phi^2(x, t) \frac{\partial^2 g(x, t)}{\partial x^2} = 0 \quad (13)$$

with the terminal condition $g(x, t) = h(x)$ for all x , where we assume that the stochastic process $g(X(t), t)$, $0 \leq t \leq T$ is a martingale.^{7,8}

From Eq. (3), we obtain the time development of an arbitrary $f(X(t))$ using the Itô formula⁶

$$f[X(t) + dX(t)] - f[X(t)] = \frac{\partial f}{\partial x} \left[K(X(t), t) dt + [P(X(t), t)]^{\frac{1-q}{2}} \circ dW \right] + \frac{1}{2} \frac{\partial^2 f}{\partial x^2} \left\{ [P(X(t), t)]^{\frac{1-q}{2}} \right\} + \mathcal{O}(\dots). \quad (14)$$

Taking the average of both sides in the equation above, we obtain

$$\left\langle \frac{df}{dt} \right\rangle = \left\langle \left[\frac{df}{dx} K(X(t), t) + \frac{1}{2} \frac{\partial^2 f}{\partial x^2} \left\{ [P(X(t), t)]^{\frac{1-q}{2}} \right\} \right] \right\rangle + \frac{d}{dt} \left\langle \left\{ [P(X(t), t)]^{\frac{1-q}{2}} \right\} \right\rangle \quad (15)$$

and using

$$\begin{aligned} \frac{\langle f(X(t)) \rangle}{dt} &= \frac{d}{dt} \int_{-\infty}^{\infty} dx f(x) P(x, t) \\ &= \int_{-\infty}^{\infty} dx f(x) \frac{\partial}{\partial t} [P(x, t)] = \int_{-\infty}^{\infty} \frac{\partial f}{\partial x} K(x, t) P(x, t) dx + \frac{1}{2} \int_{-\infty}^{\infty} \frac{\partial^2 f}{\partial x^2} \left\{ [P(x, t)]^{\frac{1-q}{2}} \right\} P(x, t) dx + \int_{-\infty}^{\infty} dx f(x) \frac{\partial P}{\partial t}. \end{aligned} \quad (16)$$

We integrate by parts and discard surface terms to obtain

$$\int_{-\infty}^{\infty} dx f(x) \frac{\partial}{\partial t} [P(x, t)] = \int_{-\infty}^{\infty} f(x) \frac{\partial}{\partial x} [K(x, t) P(x, t)] dx + \frac{1}{2} \int_{-\infty}^{\infty} f(x) \frac{\partial^2}{\partial x^2} \left[\left\{ [P(x, t)]^{\frac{1-q}{2}} \right\} P(x, t) \right] dx. \quad (17)$$

and therefore

$$\frac{\partial}{\partial t} P(x, t) = - \frac{\partial}{\partial x} [K(x, t) P(x, t)] + \frac{1}{2} \frac{\partial^2}{\partial x^2} \left\{ \left\{ [P(x, t)]^{\frac{1-q}{2}} \right\}^2 P(x, t) \right\}. \quad (18)$$

Thus, we have a complete equivalence between the diffusion process defined by a drift coefficient $K(x, t)$ and a diffusion coefficient given as $\phi(x, t) = [P(x, t)]^{1-q}$, in which the diffusion process can be locally approximated by an Itô stochastic differential equation.

Correspondence to Stratonovich stochastic differential equation

$$(S)dX = K^s(X(t), t)dt + [P(X(t), t)]^{\frac{1-q}{2}} dW, \quad (19)$$

with $K^s = K - \frac{1}{2}\phi\partial_x\phi$ and using the correspondence between the Itô stochastic differential equation and the Fokker-Planck equation, we have a equivalent Fokker-Planck equation

$$\partial_t P = -\partial_x \{K^s P\} + \frac{1}{2}\partial_x \{\phi\partial_x [\phi P]\}, \quad (20)$$

which is known as Stratonovich form of the Fokker-Planck equation⁶. However, it is different from Eq. (2). Therefore, we have that the corresponding nonlinear Fokker-Planck equation in the Stratonovich prescription is different from nonlinear equation obtained in the Itô prescription, being the Itô stochastic differential equation more usually employed to make the connection with the Fokker-Planck equation, contrary to Ref.¹⁹. In spite of both definitions can be related by the choosing of i by $\tau_i = \alpha t_i + (1 - \alpha)t_{i-1}$, $\alpha \in (0, 1)$ they generate different definitions of stochastic integral (Itô integral and Stratonovich integral) and consequently to different stochastic differential equations, even though the Itô stochastic differential equation is equivalent to an another Stratonovich equation however, with an additional term⁶.

2.2 Existence and Uniqueness

We can investigate the existence and uniqueness of solutions of nonlinear differential equations utilizing the well-known existence and uniqueness theorem for stochastic differential equations⁷. Let $T > 0$ and $K(x, t) : [0, T] \times \mathbb{R}^n \rightarrow \mathbb{R}^n$, $\phi(x, t) : [0, T] \times \mathbb{R}^n \rightarrow \mathbb{R}^{n \times m}$ be measurable functions satisfying

$$|K(x, t)| + |\phi(x, t)| \leq C(1 + |x|); x \in \mathbb{R}^n, \quad t \in [0, T] \quad (21)$$

for some constant C and such that

$$|b(x, t) - b(y, t)| + |(\phi(x, t) - \phi(y, t))| \leq D|x - y|; \quad x, y \in \mathbb{R}^n, \quad t \in [0, T] \quad (22)$$

for some constant D . Let Z be a random variable which is independent on the σ -algebra \mathcal{F}_∞^m generated by $W_s(\cdot)$, $s \leq 0$ and such that the expectation $E[|Z|^2] < \infty$. Then the stochastic differential equation

$$dX = A(X(t), t)dt + \phi(X(t), t)dW, \quad 0 \leq t \leq T, \quad X_0 = Z \quad (23)$$

has a unique t -continuous solution $X_t(\omega)$ with the property that $X_t(\omega)$ is adapted to filtration \mathcal{F}_t^Z generated by Z and $W_s(\cdot)$; $s \leq t$

$$E \left[\int_0^T |X_t|^2 dt \right] < \infty. \quad (24)$$

3 Analysis by Fokker-Planck equation

3.1 Stochastic model for novel cases

The behavior of the novel cases number $N(t)$ of infected by coronavirus registered in the Brazil as function of time (days) registered since 15th March, 2020 is displayed in Fig. 2. From fit of least minimums squares to the set of official data of the Brazilian ministry of health, we estimate the curve for the behavior of the novel cases of COVID-19 given by a polynomial of fourth order of t (solid-blue-line) in the Fig. 2 given as $N(t) = -1152.5 + 126.1t - 4.0925t^2 + 0.064749t^3 - 2.6 \times 10^{-5}t^4$. The data are by May 22th. The zig-zag behavior into the range of t considered reflects in an increase of the uncertainty in the data and to population isolation conditions. For modeling of the situation, we added a random term together with nonlinear terms in Eq. (1) with aim to simulate the effect of this uncertainty. The α and ν parameters and the nonlinear polynomial added comes from the Bernoulli model for the spreading of the smallpox with the value $\alpha = \nu = 1/8$ obtained by him. As the COVID-19 as the smallpox, once contracted and survived, confers a lifetime immunity, considering a group of individuals in a given time t where we have a number of these individuals surviving a time t later. We propose then that the model for the novel cases in each day t obeys to the following stochastic model

$$dN(t) = [f(t) + A(N(t), t)]dt + B(N(t), t) \circ dW, \quad (25)$$

where $f(t)$ is a polynomial of n degree ($n = 4$). With a drift term in the form of a standard Bernoulli model of disease spread and is driven by a Wiener process with normal distribution. The deterministic part can adjust to the reported novel cases, in the average.

3.2 Analytical Results

We can perform the simulation of the model Eq. (25) with term $\beta W(t)$, Winner increment and an additive white noise of standard deviation $\sigma_w = \sqrt{\Delta t}$. We can write the Winner increment as $\beta dW(t) \sim \sqrt{dt}\beta R_G$, where R_G is an aleatory generator number with Gaussian distribution of mean zero and variance $\sigma_w^2 = 1$.

From stochastic equation Eq. (25), we obtain the time development of an arbitrary function $f(X(t))$ by using of the Itô formula

$$f[X(t) + dX(t)] - f[X(t)] = \partial_x f[X(t)] \{ [f(t') - \alpha X(t')(1 - vX(t'))] dt + \beta dW \} + \frac{\beta^2}{2} \partial_x^2 f[X(t)] (dW)^2, \quad (26)$$

where higher order terms have been discarded, and $(dW(t))^2 = dt$. Taking the average of both sides in the equation above and defining $\gamma = \beta^2$, (β constant) we obtain

$$\left\langle \frac{\partial f}{\partial t} \right\rangle = \left\langle \left[\frac{\partial f}{\partial x} \{ [f(t) - \alpha x(1 - vx)] dt + \beta dW \} + \frac{\gamma}{2} \frac{\partial^2 f}{\partial x^2} \right] \right\rangle. \quad (27)$$

Using

$$\begin{aligned} \frac{d}{dt} \langle f(X(t)) \rangle &= \frac{d}{dt} \int_{-\infty}^{\infty} dx f(x) P(x, t) = \int_{-\infty}^{\infty} dx f(x) \frac{\partial}{\partial t} [P(x, t)] = \int_{-\infty}^{\infty} \frac{\partial f}{\partial x} \{ [f(t) - \alpha x(1 - vx)] \} P(x, t) dx \\ &\quad + \frac{\gamma}{2} \int_{-\infty}^{\infty} \frac{\partial^2 f}{\partial x^2} P(x, t) dx. \end{aligned} \quad (28)$$

We integrate by parts and discard surface terms to obtain

$$\int_{-\infty}^{\infty} dx f(x) \frac{\partial}{\partial t} [P(x, t)] = \int_{-\infty}^{\infty} f(x) \frac{\partial}{\partial x} \{ [f(t) - \alpha x(1 - vx)] P(x, t) \} dx + \frac{\gamma}{2} \int_{-\infty}^{\infty} f(x) \frac{\partial^2}{\partial x^2} [P(x, t)] dx. \quad (29)$$

and hence

$$\frac{\partial}{\partial t} P(x, t) = - \frac{\partial}{\partial x} \{ [f(t) - \alpha x(1 - vx)] P(x, t) \} + \frac{\gamma}{2} \frac{\partial^2}{\partial x^2} P(x, t). \quad (30)$$

The associated Fokker-Planck equation to the above model is given by

$$\frac{\partial P(x, t)}{\partial t} = \frac{\partial}{\partial x} \{ [f(t) - \alpha x(1 - vx)] P(x, t) \} + \frac{\gamma}{2} \frac{\partial^2 [P(x, t)]}{\partial x^2}. \quad (31)$$

taking the Fourier transform of the Fokker-Planck equation, we can guarantee the normalization of the probability density in which $P(x)$ is reasonably well behaved. We take the boundaries at infinity for $P(x, t)$ as $\lim_{x \rightarrow \infty} P(x, t) = 0$ and therefore $\partial_x P(x)$ being reasonably well behaved. As $\lim_{x \rightarrow \infty} \partial_x P(x, t) = 0$ thus, a nonzero current of probability at infinity will usually require that the terms in the equation above become infinite there⁶. We use the initial condition $P(x = x_0, t = 0) = P_0$.

For solving the Fokker-Planck equation above independent on time we make the power series expansion $P(x, t) = \sum_{n=0}^{\infty} a_n(t) x^n$ to obtain

$$\begin{aligned} \frac{\partial P}{\partial t} &= \alpha(1 - 2vx)P - [f(t) - \alpha x(1 - vx)] \frac{\partial P}{\partial x} + \frac{\gamma}{2} \frac{\partial^2 P}{\partial x^2} \\ \sum_{n=0}^{\infty} \left(\frac{da_n}{dt} + f(t)(n+1)a_{n+1} \right) x^n &= [\alpha(1 - vx) - \alpha vx] \sum_{n=0}^{\infty} a_n x^n + \alpha x(1 - vx) \sum_{n=0}^{\infty} n a_n x^{n-1} + \frac{\gamma}{2} \sum_{n=0}^{\infty} n(n-1) a_n x^{n-2}. \end{aligned} \quad (32)$$

We obtain the relations

$$\alpha \sum_{n=0}^{\infty} a_n x^n - 2v\alpha \sum_{n=1}^{\infty} a_{n-1} x^n + \sum_{n=1}^{\infty} n a_n x^n - \alpha v \sum_{n=2}^{\infty} (n-1) a_{n-1} x^n + \frac{\gamma}{2} \sum_{n=0}^{\infty} (n+1)(n+2) a_{n+2} x^n = k, \quad (33)$$

where k is a separation constant. For $k = 0$, one obtains the following recurrence relations

$$\begin{aligned} a_2 &= -\frac{\alpha}{\gamma} a_0 \\ a_3 &= 2v\alpha a_0 - a_1(1 + \alpha) \\ a_4 &= \frac{\alpha}{3\gamma^2} \left(\frac{\alpha}{2} - 1 \right) a_0 + \frac{v\alpha}{2\gamma} a_1 \end{aligned} \quad (34)$$

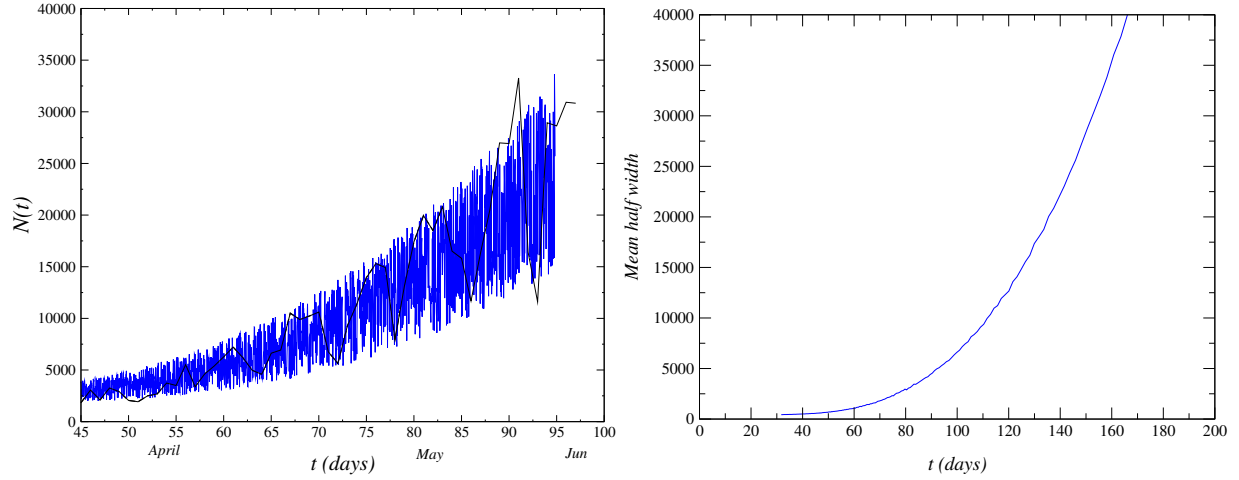


Figure 3. Time evolution of new cases (left) and behaviour of the mean half- width (right) of the distribution σ with t . The half-width gives an expectation of new cases in the day t .

for $0 \leq n \leq 2$, and

$$a_{n+2} = \frac{2}{\gamma} \frac{[\alpha v(n+1)a_{n-1} - (\alpha + n)a_n]}{(n+2)(n+1)}, \quad (35)$$

for $n > 2$. Additionally, we have

$$\frac{da_n(t)}{dt} + f(t)(n+1)a_{n+1}(t) = k. \quad (36)$$

Therefore, we obtain $P(x)$ in the form

$$P(x) = \left\{ 1 - \frac{\alpha}{\gamma}x^2 + 2v\alpha x^3 + \frac{\alpha}{3\gamma^2} \left(\frac{\alpha}{2} - 1 \right) x^4 - \frac{1}{10\gamma} \left[2v(3+\alpha) + \frac{2v\alpha}{\gamma}(v+\alpha) \right] x^5 \dots \right\} a_0 \\ + \left\{ x - (1+\alpha)x^3 + \frac{v\alpha}{2\gamma}x^4 + \frac{1}{10\gamma}(3+\alpha)(1+\alpha)x^5 + \dots \right\} a_1. \quad (37)$$

The constants a_0 and a_1 are determined by the initial conditions $P(0,0) = P_0$ and $\partial_x P(x,0) = 0$ in $x = 0$. We obtain $a_0 = P_0$ and $a_1 = 0$. From the normalization condition, the second term in the density probability above must be zero and therefore, all coefficients a_1 must cancel. Therefore, we have

$$P(x,t) = P_0 \left\{ 1 - \frac{\alpha}{\gamma}x^2 + 2v\alpha x^3 + \frac{\alpha}{3\gamma^2} \left(\frac{\alpha}{2} - 1 \right) x^4 - \frac{1}{10\gamma} \left[2v(3+\alpha) + \frac{2v\alpha}{\gamma}(v+\alpha) \right] x^5 \dots \right\} \quad (38)$$

To ensure the normalization of the probability density, P_0 must be non zero only within interval $-\varepsilon \leq x \leq \varepsilon$ and zero out it.

For $k \neq 0$, we have from Eq. (33) that $n = 0$ and $a_2 + \alpha a_0/\gamma = k$ and all a_n higher are zero. Consequently, we obtain from integration of the Eq. (35)

$$P(t) = \frac{\gamma}{\alpha} P_0 t [1 - a_2(t)] + a_0(0) - \int_0^t f(t') a_1(t') dt'. \quad (39)$$

We obtain the mean half-width of the distribution as function of time as given as

$$\sigma(t) = \left\{ \frac{[N(t)]^3}{3} \frac{\gamma}{\alpha} \left[P_0(1 - a_2(t)) + a_n(0) - \int_0^t f(t') a_1(t') dt' \right]^2 - \frac{[N(t)]^4}{4} \frac{\gamma^2}{\alpha^2} \left[P_0(1 - a_2(t)) + a_n(0) - \int_0^t f(t') a_1(t') dt' \right]^2 \right\}^{1/2}$$

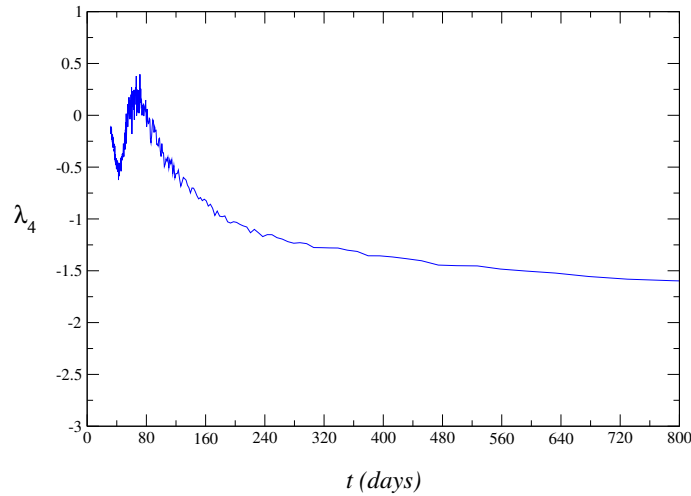


Figure 4. Time evolution of the kurtosis $\lambda_4(t)$. The negative values obtained for all t values gives an estimating of the distribution shape that becomes nearest of a Gaussian $\lambda_4 = 0$ at range of large t values from firsts cases registered in Brazil from 15th March 2020.

Furthermore, we have $(n = 0) da_0(t)/dt = 0$, or $a_0(t) = c$, $a_1 = k/2f(t)$ and $P(t) = P_0t + c - k/2$, where we define $p_0 = c - k/2$. Therefore

$$\sigma(t) = (P_0t + p_0) \sqrt{\frac{[N(t)]^3}{3} - \frac{[N(t)]^4}{4}}. \quad (40)$$

How the first cases are registered on March 15th ($t = 15$), we obtain $p_0 = -1.5$ and $P_0 = 0.1$ for the official results of the novel cases number.

In Fig. 3(left), we plot the time series of novel cases registered in the Brazil and the time series of the model Eq. (25). We obtain that the model adjust well to the data within range considered. In right plot we obtain the mean half-width as function of time t (days). We calculate the variance of the distribution of novel cases, where the standard deviation gives an estimating of the number of the novel cases at day t . The results fit qualitatively to the official data of ministry of healthy. According of the results obtained we obtain a forecasting of growing in the new cases number in the next weeks.

In Fig. 4, we display the behavior of the kurtosis (excess), $\lambda_4(t)$ as function of time. The excess kurtosis relates to the deviation of the tail of the distribution, as compared to Gaussian. The negative value obtained for the kurtosis for large t values indicates that the shape of the distribution is near to the Wigner semicircle. Furthermore, at range of small t where the kurtosis is nearest to zero we have that the distribution is nearest of a Gaussian ($\lambda_4 = 0$). The small oscillation in the range short time from the beginning of the registration of cases should be due to the low number of data in this range of time.

4 Conclusions

In Brief, we propose a stochastic model for the time evolution of the SARS-CoV-2 (COVID-19) in Brazil based in the nonlinear Itô's diffusion. Our results are compared with official data supplied by the Brazilian healthy agencies where due to large uncertain in the results generated principally by the low number of tests made in the population and hence, to under reporting, generates a large uncertainly in the official results and consequently, the stochastic differential equation analysis becomes the more realistic model for growing of the total number of infected \mathcal{P} and novel cases number $N(t)$. We can use an approach beyond white noise limit to try to better describe the expansion dynamics what can be done in a future work. The model reported here is based on Brazilian data from March 15 which shows an upward trend in the coming weeks.

References

1. Strogatz, S. H. Nonlinear dynamics and chaos (1994).
2. W. E Boyce, R. C. D. Elementary differential equations (2012).
3. Habib, G., G. and Kerschen. A principle of similarity for nonlinear vibration absorbers. *Phys. D: Nonlinear Phenom.* **332**, 1–8, DOI: <https://doi.org/10.1016/j.physd.2016.06.001> (2016).

4. Fernandez-Garcia, S., Krupap, M. & Clement, F. Mixed-mode oscillations in a piecewise linear system with multiple time scale coupling. *Phys. D: Nonlinear Phenom.* **332**, 9–22, DOI: <https://doi.org/10.1016/j.physd.2016.06.002> (2016).
5. Nobre, F. D., Rego-Monteiro, M. A. & C., T. A generalized nonlinear schrödinger equation: Classical field-theoretic approach. *EPL (Europhysics Lett.* **97**, 41001, DOI: <https://doi.org/10.12092F0295-50752F972F41001> (2012).
6. Gardiner, C. Stochastic methods, a handbook for the natural and social sciences (2009).
7. Øksendal, B. Stochastic differential equations an introduction with applications (2013).
8. Steven, E. S. Stochastic calculus for finance ii continuous-time models (2004).
9. Mantegna, R. S. & Stanley, H. E. An introduction to econophysics, correlations and complexity in finance (2012).
10. Gopikrishnan, P., Meyer, M., Amaral, L. A. N. H. & Stanley, H. E. Inverse cubic law for the distribution of stock price variations. *Eur. Phys. J. B* **3**, 139–140, DOI: <https://doi.org/10.1007/s100510050292> (1998).
11. Gopikrishnan, P., Plerou, V., Amaral, L. A. N. H., Meyer, M. & Stanley, H. E. Scaling of the distribution of fluctuations of financial market indices. *Phys. Rev. E* **60**, 5305, DOI: <https://doi.org/10.1103/PhysRevE.60.5305> (1999).
12. Plerou, V. *et al.* Scaling of the distribution of price fluctuations of individual companies. *Phys. Rev. E* **60**, 6519, DOI: <https://doi.org/10.1103/PhysRevE.60.6519> (1999).
13. Botta, H. S., Fand Moat, Stanley, H. E. & T., P. Quantifying stock return distributions in financial markets. *PLoS ONE* **10**, e0135600, DOI: <https://doi.org/10.1371/journal.pone.0135600> (2015).
14. Gu, G.-F. Z. & W.-X. Pagerank equation and localization in the www. *EPL (Europhysics Lett.* **86**, 48002, DOI: <https://doi.org/10.1209/0295-5075/88/48002> (2009).
15. Gu, G.-F. & Zhou, W.-X. On the probability distribution of stock returns in the mike-farmer model. *Eur. Phys. J. B* **67**, 585–592, DOI: <https://doi.org/10.1140/epjb/e2009-00052-4> (2009).
16. Meng, H. F. *et al.* On the probability distribution of stock returns in the mike-farmer model. *EPL (Europhysics Lett.* **98**, 38003, DOI: <https://doi.org/10.1209/0295-5075/98/38003> (2012).
17. Meng, H. F. *et al.* Computational experiments successfully predict the emergence of autocorrelations in ultra-high-frequency stock returns. *Comput. Econ.* **50**, 579–594, DOI: <https://doi.org/10.1007/s10614-016-9612-1> (2017).
18. Plastino, A. R., Curado, E. M. F., Nobre, F. D. & Tsallis, C. From the nonlinear fokker-planck equation to the vlasov description and back: Confined interacting particles with drag. *Phys. Rev. E* **97**, 022120, DOI: <https://doi.org/10.1103/PhysRevE.97.022120> (2018).
19. Arenas, Z. G., Barci, D. G. & Tsallis, C. Nonlinear inhomogeneous fokker-planck equation within a generalized stratonovich prescription. *Phys. Rev. E* **90**, 032118, DOI: <https://doi.org/10.1103/PhysRevE.90.032118> (2014).

Acknowledgements

This work was partially supported by the Brazilian agency CNPq.

Author contributions statement

Leonardo S. Lima is the sole author of this manuscript. He has conducted the experiment, analysed the results and written the paper.

Figures

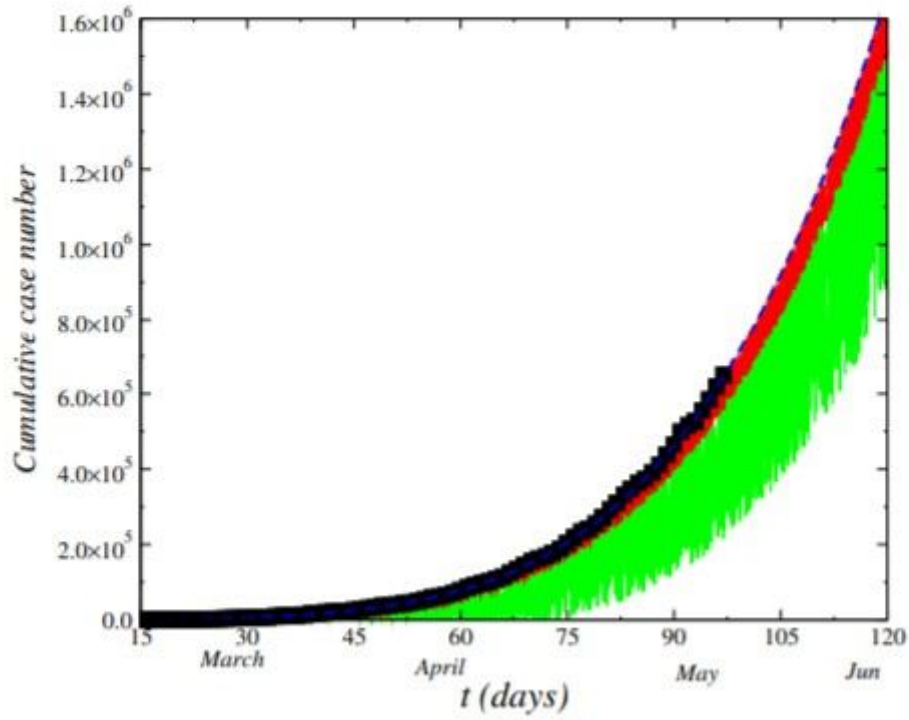


Figure 1

Effect of the stochastic term on cumulative total case number of the COVID-19 in Brazil. We perform the calculations for different values of β_0 in model Eq. (1) as $\beta_0 = 0.0$ (dashed blue-line), $\beta_0 = 3.0 \times 10^{-6}$ (red-line) and $\beta_0 = 5.0 \times 10^{-6}$ (green-line).

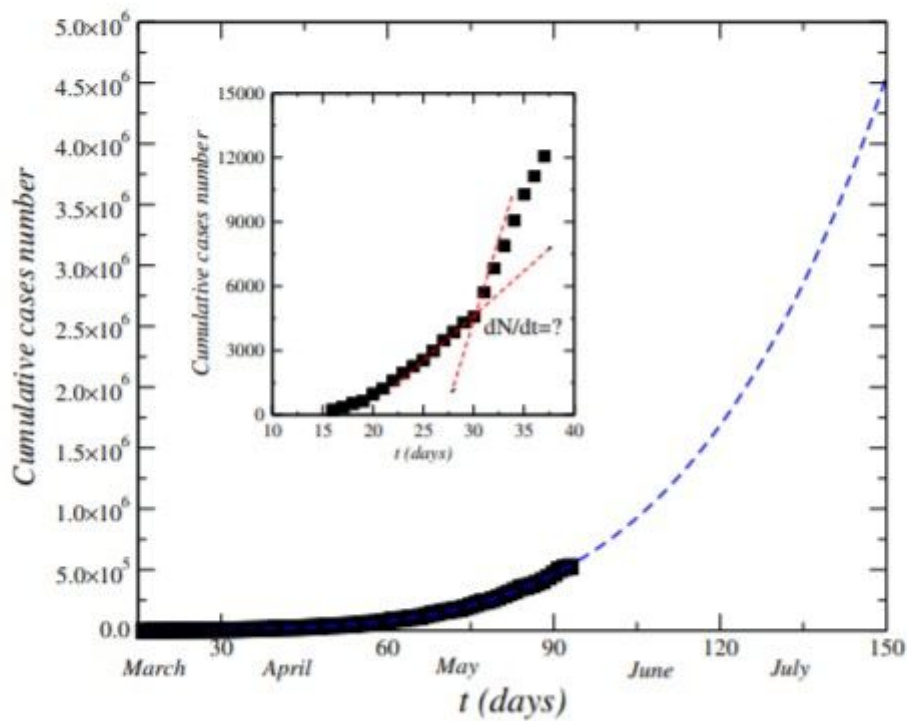


Figure 2

Total novel cases number fit of COVID-19 $P(t)$. The squares (in black) are the cumulative case number registered in the Brazil in the range considered. We obtain a polynomial adjustment in the range considered (dashed-blue-line). In the inset, we display the non differentiability of the data curve.

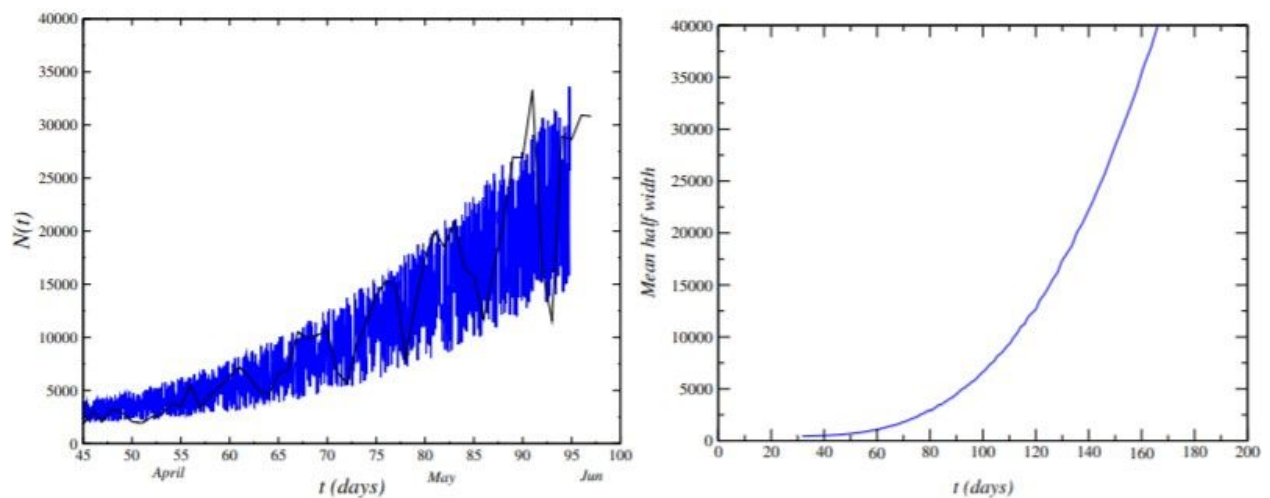


Figure 3

Time evolution of new cases (left) and behaviour of the mean half- width (right) of the distribution σ with t . The half-width gives an expectation of new cases in the day t .

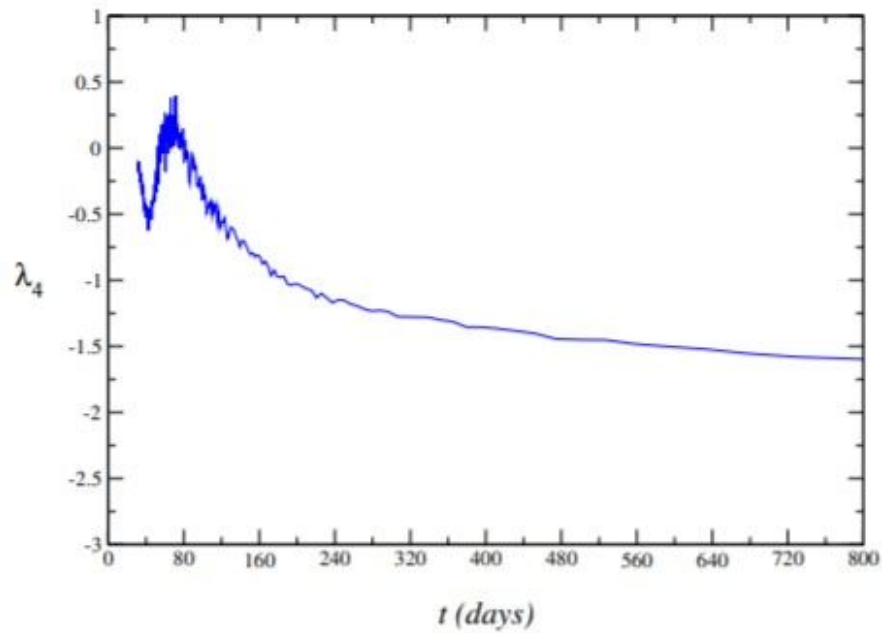


Figure 4

Time evolution of the kurtosis $\lambda_4(t)$. The negative values obtained for all t values gives an estimating of the distribution shape that becomes nearest of a Gaussian $\lambda_4 = 0$ at range of large t values from firsts cases registered in Brazil from 15th March 2020.