



# **UNIVERSIDAD DE INVESTIGACIÓN DE TECNOLOGÍA EXPERIMENTAL YACHAY**

**Escuela de Ciencias Físicas y Nanotecnología**

## **TÍTULO: MODELING THE SPREAD OF COVID-19 USING THE LOGISTIC MODEL**

Trabajo de integración curricular presentado como  
requisito para la obtención  
del título de Ingeniero en Nanotecnología

**Autor:**

Cachago Pillajo Dennis Paul

**Tutor:**

Ph.D. Vacacela Gomez Cristian Isaac

Urcuquí, abril de 2022

**SECRETARÍA GENERAL**  
(Vicerrectorado Académico/Cancillería)  
**ESCUELA DE CIENCIAS FÍSICAS Y NANOTECNOLOGÍA**  
**CARRERA DE NANOTECNOLOGÍA**  
**ACTA DE DEFENSA No. UITEY-PHY-2022-00004-AD**

A los 31 días del mes de marzo de 2022, a las 14:00 horas, de manera virtual mediante videoconferencia, y ante el Tribunal Calificador, integrado por los docentes:

<b>Presidente Tribunal de Defensa</b>	Dr. COSENZA MICELI, MARIO GIUSEPPE , Ph.D.
<b>Miembro No Tutor</b>	Dr. GUEVARA GRANIZO, MARCO VINICIO , Ph.D.
<b>Tutor</b>	Dr. VACACELA GOMEZ, CRISTIAN ISAAC , Ph.D.

El(la) señor(ita) estudiante **CACHAGO PILLAJO, DENNIS PAUL**, con cédula de identidad No. **1718423633**, de la **ESCUELA DE CIENCIAS FÍSICAS Y NANOTECNOLOGÍA**, de la Carrera de **NANOTECNOLOGÍA**, aprobada por el Consejo de Educación Superior (CES), mediante Resolución **RPC-SO-13-No.156-2015**, realiza a través de videoconferencia, la sustentación de su trabajo de titulación denominado: **MODELING THE SPREAD OF COVID-19 USING THE LOGISTIC MODEL**, previa a la obtención del título de **INGENIERO/A EN NANOTECNOLOGÍA**.

El citado trabajo de titulación, fue debidamente aprobado por el(los) docente(s):

<b>Tutor</b>	Dr. VACACELA GOMEZ, CRISTIAN ISAAC , Ph.D.
--------------	--

Y recibió las observaciones de los otros miembros del Tribunal Calificador, las mismas que han sido incorporadas por el(la) estudiante.

Previamente cumplidos los requisitos legales y reglamentarios, el trabajo de titulación fue sustentado por el(la) estudiante y examinado por los miembros del Tribunal Calificador. Escuchada la sustentación del trabajo de titulación a través de videoconferencia, que integró la exposición de el(la) estudiante sobre el contenido de la misma y las preguntas formuladas por los miembros del Tribunal, se califica la sustentación del trabajo de titulación con las siguientes calificaciones:

Tipo	Docente	Calificación
Presidente Tribunal De Defensa	Dr. COSENZA MICELI, MARIO GIUSEPPE , Ph.D.	10,0
Tutor	Dr. VACACELA GOMEZ, CRISTIAN ISAAC , Ph.D.	10,0
Miembro Tribunal De Defensa	Dr. GUEVARA GRANIZO, MARCO VINICIO , Ph.D.	9,4

Lo que da un promedio de: **9.8 (Nueve punto Ocho)**, sobre 10 (diez), equivalente a: **APROBADO**

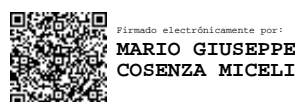
Para constancia de lo actuado, firman los miembros del Tribunal Calificador, el/la estudiante y el/la secretario ad-hoc.

Certifico que *en cumplimiento del Decreto Ejecutivo 1017 de 16 de marzo de 2020, la defensa de trabajo de titulación (o examen de grado modalidad teórico práctica) se realizó vía virtual, por lo que las firmas de los miembros del Tribunal de Defensa de Grado, constan en forma digital.*

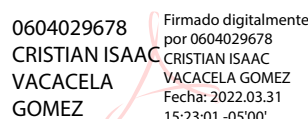
**CACHAGO PILLAJO, DENNIS PAUL**  
**Estudiante**



**Dr. COSENZA MICELI, MARIO GIUSEPPE , Ph.D.**  
**Presidente Tribunal de Defensa**



**Dr. VACACELA GOMEZ, CRISTIAN ISAAC , Ph.D.**  
**Tutor**



Dr. GUEVARA GRANIZO, MARCO VINICIO , Ph.D.  
**Miembro No Tutor**

MARCO  
VINICIO  
GUEVARA  
GRANIZO

Firmado digitalmente  
por MARCO VINICIO  
GUEVARA GRANIZO  
Fecha: 2022.03.31  
15:25:10 -05'00'

CIFUENTES TAFUR, EVELYN CAROLINA  
**Secretario Ad-hoc**

EVELYN  
CAROLINA  
CIFUENTES TAFUR

Firmado digitalmente por  
EVELYN CAROLINA CIFUENTES  
TAFUR  
Fecha: 2022.03.31 14:50:08  
-05'00'

## **AUTORÍA**

Yo, **DENNIS PAUL CACHAGO PILLAJO**, con cédula de identidad 1718423633, declaro que las ideas, juicios, valoraciones, interpretaciones, consultas bibliográficas, definiciones y conceptualizaciones expuestas en el presente trabajo; así cómo, los procedimientos y herramientas utilizadas en la investigación, son de absoluta responsabilidad de el/la autora (a) del trabajo de integración curricular. Así mismo, me acojo a los reglamentos internos de la Universidad de Investigación de Tecnología Experimental Yachay.

Urcuquí, abril de 2022

---

Dennis Paul Cachago Pillajo

CI: 1718423633

## **AUTORIZACIÓN DE PUBLICACIÓN**

Yo, **DENNIS PAUL CACHAGO PILLAJO**, con cédula de identidad 1718423633, cedo a la Universidad de Investigación de Tecnología Experimental Yachay, los derechos de publicación de la presente obra, sin que deba haber un reconocimiento económico por este concepto. Declaro además que el texto del presente trabajo de titulación no podrá ser cedido a ninguna empresa editorial para su publicación u otros fines, sin contar previamente con la autorización escrita de la Universidad.

Asimismo, autorizo a la Universidad que realice la digitalización y publicación de este trabajo de integración curricular en el repositorio virtual, de conformidad a lo dispuesto en el Art. 144 de la Ley Orgánica de Educación Superior

Urcuquí, abril de 2022.

---

Dennis Paul Cachago Pillajo

CI: 1718423633

## **Acknowledgements**

First of all, I want to say thank you very much to my advisor, Cristian Vacacela. Thank you for receiving me as your student and giving me the opportunity to carry out this research work with you. For keeping me motivated during the trip. For being patient all the times that I did not understand. For believing in me. The end of this adventure would not have been possible without your help.

To my parents Yolanda and Luis for all the love, support and patience they have given me. To my sister, the person who I love most in the whole world and for whom I would give everything to see her happy. I know that with this, I will. To my brother, who inspired me to not stop doing what I love, and who teach me that no one can judge me than myself. A special thanks, to my soulmate, for always motivating me to be a better person, a better student and a better son.

I also want to thank the School of Physical Sciences and Nanotechnology for the quality of teachers that has, because they motivated us to do not stop our dreams and follow the path of science. Henry, Leonardo, Jorge, Cristian, Sarah, Gema, Werner, and the other teachers, thank you a lot. To my friends, inside and outside the university who have trusted and have supported me during this adventure.

Thank you all.

## Resumen

La situación actual debido a la pandemia de COVID-19 está involucrando rápidamente a investigadores de todo el mundo que se dedican a promover modelos matemáticos adecuados para comprender la propagación de esta enfermedad. Sin embargo, la gran cantidad de datos y los diferentes patrones dificultan mantenerse al día con la propagación y el desarrollo de COVID-19. El presente trabajo propone un enfoque matemático para explorar la propagación del COVID-19 en América Latina. Por su simplicidad, este modelo matemático se puede utilizar para estimar, principalmente, la tasa de crecimiento de la infección, el punto de inflexión y el número máximo de casos. El enfoque propuesto se basa en la ecuación logística discreta y diferencial, así como el modelo exponencial se utilizan para escudriñar el efecto de COVID-19 al comienzo de la pandemia. Se utilizó un conjunto de datos de un año para diferentes países con un intervalo de confianza del 95% y una media móvil de siete días para mejorar la resolución de los datos recopilados. Finalmente, este enfoque matemático se puede extender para analizar el efecto de COVID-19 y predecir sus consecuencias en otros lugares, lo que permite tomar nuevas decisiones frente a la enfermedad de COVID-19.

**Palabras clave:** COVID-19, modelo logístico, Sur América, modelo matemático, propagación en multi etapas.

## Abstract

The current situation due to the COVID-19 pandemic is rapidly involving researchers from around the world who are dedicated to promote suitable mathematical models to understand the spread of this disease. However, the large amount of data and different patterns make it difficult to keep up with the spread and development of COVID-19. The present work proposes a mathematical approach to explore the spread of COVID-19 in Latin America. Because of its simplicity, this mathematical model can be used to estimate, mainly, the growth infection rate, turning point, and maximum number of cases. The proposed approach is based on the discrete and differential logistic equation, as well as, the exponential model are used to scrutinize the effect of COVID-19 at the beginning of the pandemic. A one-year data set was used for different countries with a confidence interval of 95% and a moving-average of seven days to improve the collected-data resolution. Finally, this mathematical approach can be extended to analyze the effect of COVID-19 and to predict its consequences in other locations, allowing revenue new decisions against the COVID-19 disease.

**Keywords:** COVID-19, logistic model, South America, mathematical model, multi-stage spread





# Contents

<b>List of Figures</b>	<b>x</b>
<b>List of Tables</b>	<b>xiv</b>
<b>1 Introduction</b>	<b>1</b>
1.1 Pandemic outbreak in South America.....	2
1.2 Mathematical Model.....	3
1.3 General Classification of Mathematical Models.....	5
1.4 Epidemic Mathematical Models .....	5
1.5 Models .....	7
1.5.1 Exponential Model.....	7
1.5.2 Logistic Model .....	7
1.6 General and Specific Objectives .....	10
1.6.1 General Objective.....	10
1.6.2 Specific Objectives.....	10
<b>2 Methodology</b>	<b>11</b>
2.1 Theoretical Approach.....	11
2.2 Data Collection and Analysis.....	15
<b>3 Results &amp; Discussion</b>	<b>17</b>
3.1 South America daily and cumulative cases.....	17
3.2 Exponential Model.....	24
3.3 Discrete Logistic Model.....	28

3.4	Differential Logistic Model .....	33
3.5	Differential Logistic Model in Function of Daily and Accumulated Cases.....	37
3.6	COVID-19 Outbreak in Ecuador .....	41
<b>4</b>	<b>Conclusions &amp; Outlook</b>	<b>55</b>
<b>A</b>	<b>Short Appendix 1 Heading for the Table of Contents</b>	<b>57</b>
A.1	Exponential function calculus.....	57
A.2	Solution for ODE differential logistic equation.....	58
	<b>Bibliography</b>	<b>61</b>

# List of Figures

2.1	Illustration cases: a) the total number of cases as a function of time using the exponential model, b) the total number of cases as a function of time. The discrete logistic equations (red line) are compared to the solution of the ODE logistic model (dashed blue line), and c) the daily cases as a function of total cases. Parameters used: $N_{max} = 100$ , $N_0 = 1$ , $\lambda_{log} = 0.2(day - 1)$ , $\lambda_{exp} = 0.2(day - 1)$ , $t_0 = 30 (day)$ .....	12
2.2	Schematic representation of the research methodology.....	14
3.1	Cumulative cases map of confirmed COVID-19 cases until Feb 28, 2021. Source: COVID-19 Data Repository by the Center for Systems Science and Engineering (CSSE) at Johns Hopkins University ( <a href="https://github.com/CSSEGISandData/COVID-19">https://github.com/CSSEGISandData/COVID-19</a> ). .....	18
3.2	a) the number of infected people per day in Argentina, b) the total number of confirmed cases reported during 365 days in Argentina, c) the number of infected people per day in Brazil, d) the total number of confirmed cases reported during 365 days in Brazil, e) the number of infected people per day in Bolivia, f) the total number of confirmed cases reported during 365 days in Bolivia. Data are represented in black and the 7-point moving average data in red. ....	20
3.3	a) the number of infected people per day in Chile, b) the total number of confirmed cases reported during 365 days in Chile, c) the number of infected people per day in Colombia, d) the total number of confirmed cases reported during 365 days in Colombia, e) the number of infected people per day in Ecuador, f) the total number of confirmed cases reported during 365 days in Ecuador. Data are represented in black and the 7-point moving average data in red. ....	21

3.4	a) the number of infected people per day in Guyana, b) the total number of confirmed cases reported during 365 days in Guyana, c) the number of infected people per day in Paraguay, d) the total number of confirmed cases reported during 365 days in Paraguay, e) the number of infected people per day in Peru, f) the total number of confirmed cases reported during 365 days in Peru. Data are represented in black and the 7-point moving average data in red. ....	22
3.5	a) the number of infected people per day in Suriname, b) the total number of confirmed cases reported during 365 days in Suriname, c) the number of infected people per day in Uruguay, d) the total number of confirmed cases reported during 365 days in Uruguay, e) the number of infected people per day in Venezuela, f) the total number of confirmed cases reported during 365 days in Venezuela. Data are represented in black and the 7-point moving average data in red.....	23
3.6	Exponential model using 30 days with 95% of confidence level. a) Argentina, b) Brazil, c) Bolivia, d) Chile, e) Colombia, f) Ecuador. Total cases data as function of time (red points) and fitting model with bands (blue lines).....	25
3.7	Exponential model using 30 days with 95% of confidence level. g) Guyana, h) Paraguay, i) Peru, j) Suriname, k)Uruguay, l) Venezuela. Total cases data as function of time (red points) and fitting model with bands (blue lines).....	26
3.8	Discrete logistic model using moving average data with 95% of confidence level. a) Argentina, b) Brazil, c) Bolivia, d) Chile, e) Colombia, f) Ecuador. Total cases data as function of time (red points) and fitting model with bands (blue lines).....	30
3.9	Discrete logistic model using data moving average with 95% of confidence level. a) Guyana, b) Paraguay, c) Peru, d) Suriname, e) Uruguay, f) Venezuela. Total cases data as function of time (red points) and fitting model with bands (blue lines). 31	
3.10	Differential logistic model using moving average data with 95% of confidence level. a) Argentina, b)Brazil, c) Bolivia, d) Chile, e) Colombia, f) Ecuador. Total cases data as function of time (red points) and fitting model with bands (blue lines) 34	
3.11	Differential logistic model using moving average data with 95% of confidence level. a) Guyana, b)Paraguay, c) Peru, d) Suriname, e) Uruguay, f) Venezuela. Total cases data as function of time(red points) and fitting model with bands (blue lines).....	35

3.12	Differential logistic model using moving average data with 95% of confidence level. Using the relationship between daily reported cases and total cases. a) Argentina, b) Brazil, c) Bolivia, d) Chile, e) Colombia, f) Ecuador. Total cases vs Daily cases data as function of time (red points) and fitting model with bands (blue lines) .....	38
3.13	Differential logistic model using moving average data with 95% of confidence level. Using the relationship between daily reported cases and total cases. a) Guyana, b) Paraguay, c) Peru, d) Suriname, e) Uruguay, f) Venezuela. Total cases vs Daily cases data as function of time (red points) and fitting model with bands (blue lines). .....	39
3.14	Ecuador population density map.....	42
3.15	Ecuador: a) the number of infected people per day, b) the total number of confirmed cases reported in a 365-days window, and c) the forward differential method applied to the daily reported cases. China: d) the total number of confirmed cases reported on 12/31/2020. Data are represented in black and the 7-point average data in red. The inset in Figure 1d shows the fit of China data using the discrete logistic model.....	43
3.16	The total number of cases as a function of time fitted with different models: a) exponential model (Equation 2.6), b) discrete logistic equation (Equation 2.7), and c) solved ODE logistic equation (Equation 2.9). Data and fit are represented in black and red, respectively. Pink lines represent the confidence level (%95) .....	45
3.17	Matrix representation of the growth rate coefficient ( $\lambda$ ) from (a) exponential model (for simplicity is taken $\lambda = 1$ ), (b) discrete logistic equation, and (c) differential logistic model. Black markers represent the infected people affected by COVID-19. ....	48
3.18	The daily cases as a function of the total confirmed cases. The black markers show the data and the red curve is the regression according to Equation 2.11. Pink lines represent the confidence level (95%). For comparison, it is analyzed different countries: a) Ecuador, b) Brazil, c) Peru, and d) Colombia. ....	50
3.19	a), c) The growth rate coefficient as a function of the total confirmed cases computed using Equation 2.12 and considering the data in a one-year window for Ecuador. b), d) the associated histograms. ....	52

3.20	The growth rate coefficient of Equation 2.12 is represented as a surface-color plot vs 30-days and 12-months. (a) the growth rate in a one-year window and (b) the growth rate from May 1, 2020, to February 28, 2021. ....	52
------	---	----

# List of Tables

3.1	Estimated growth rate parameter ( $\lambda$ ) from the exponential model for the different countries under study and curve fit ( $R^2$ ) for each one. ....	27
3.2	Estimated cases ( $N_{max}$ ) and midpoints ( $t_0$ ) parameters with their standard errors from the discrete logistic model for the different countries under study.....	29
3.3	Estimated growth rate parameter ( $\lambda$ ) from the discrete logistic model for the different countries under study and curve fit $R^2$ for each one. ....	32
3.4	Estimated cases ( $N_{max}$ ) and growth rates ( $\lambda$ ) parameters with their standard errors from the differential logistic model for the different countries under study.....	36
3.5	Estimated cases ( $N_{max}$ ) and growth rates ( $\lambda$ ) parameters with their standard errors from the differential logistic model for the different countries under study based on daily reported cases and total cases.....	40
3.6	Estimated cases ( $N_{max}$ ) and midpoints ( $t_0$ ) parameters with their standard errors from the discrete logistic model for the different countries under study.....	41
3.7	Estimated parameters using the different models of the present work. The $\lambda$ represents the estimated (exponential or logistic) growth rate, $N_{max}$ is the maximum number of infected people, and $t_0$ denotes the midpoint. ....	46
3.8	Estimated parameters using the different models of the present work. The $\lambda$ represents the estimated (exponential or logistic) growth rate, $N_{max}$ is the maximum number of infected people, and $t_0$ denotes the midpoint. ....	51





# Chapter 1

## Introduction

Through all 2020, the world's medical situation suffers and emergency due to a novel disease called COVID-19. This outbreak, technically denominated Severe Acute Respiratory Syndrome Coronavirus 2, is a zoonotic virus similar to SARS-CoV and Middle East Respiratory Syndrome-related Coronavirus (MERS-CoV)<sup>1</sup>. COVID-19 was identified firstly in China, at the end of 2019. This disease has rapidly spread across China and several nearby regions. Once it was difficult to control and isolated, China declared this situation as a public health emergency of international concern, and time after, it was recognized as a pandemic on 11 March, 2020<sup>2</sup>.

There are many theories about the origin of how this virus was transmitted to humans. One of them has gained greater prominence throughout the pandemic, whose hypothesis mentions that the virus was transmitted to humans through an intermediate host from bats. This spread would cause severe respiratory syndrome, thus achieving a strong person-to-person transmission through the air and certain objects that the carrier has had contact with<sup>3</sup>. This hypothesis has been studied by a lot of virologists who dealt with the gene sequence of novel coronavirus, after comparing coronaviruses in other creatures. Their results revealed and proved that bats and minks might be the couple probable hosts that were found to be most similar to those of the new coronavirus<sup>4</sup>.

Since initial identification, uncountable efforts were realized to attenuate the most possible spread. However, despite strict control, it became a global pandemic. This event is a large threat and a health challenge and world economies; the disease has spread to over 100 countries worldwide and counting. Until the first year of pandemic **March, 2021** approx, a total of **192,01M**

**COVID-19 cases** have been reported worldwide, with **4.13M deaths**, with an overall case fatality rate of 2.5 percent<sup>5</sup>. According to the statistics of WHO, the United States of America (US), India, Brazil, the Russian Federation, and France are the top 5 countries with the highest number of infections in the world. The degree of infection varies from country to country, and the control strategy and degree vary according to national conditions. How the global epidemic will peak or diminish is the most concerning problem. Therefore, it is of striking significance to predict the pandemic trends of infection worldwide<sup>3</sup>.

A large number of Universities around the world have investigated and developed innumerable predicting methods for the trend forecasting of COVID-19. The models based on mathematical statistics, machine learning, and deep learning have been applied to the prediction of time series of epidemic development. The mathematical models are based on studies related to infectious disease models and their propagation and real-time artificial intelligence response<sup>3</sup>. The use of these models is to create a simplified depiction of infection spread in a population. Moreover is used to understand how an infection might increase in a near future. These predictions could assist us in to making public health plans, developing vaccination programs, treatments, preventions, and interventions<sup>4</sup>.

## 1.1 Pandemic outbreak in South America

South America, the fourth largest of the world's continents is compact and roughly triangular in shape, being broad in the north and tapering to a point—Cape Horn, Chile—in the south<sup>6</sup>. This continent lies between the two large oceans, the Pacific and the Atlantic, and as such there is a great influence of these oceans on the meteorology of this landmass. SA presents features of tropical, subtropical, and extra-tropical weather and climate. An important and distinct geographical feature of the continent is the presence of a steep and narrow mountain range extending all the way from the northern tip to the southern tip along the west coast called "Cordillera de los Andes"<sup>7</sup>. Another important feature is the tropical Amazon jungle, occupying about 35% of the total continental area and 65% of the tropical area<sup>7</sup>.

South America has a total area of about 6,878,000 square miles (17,814,000 square km), with a population of 653 962 331 people according to World Meter 2020 information. South America has around 15 countries in its extension, however, for the realization of the present work, 12 of

them have been taken because they are the countries with the most infections they have had within the continent, these are Argentina, Brazil, Bolivia, Chile, Colombia, Ecuador, Guyana, Paraguay, Peru, Suriname, Uruguay, and Venezuela. The South American region was affected by the pandemic a few weeks after Europe, with the first cases of COVID-19 registered in Brazil in late February 2020. Since then, it has spread to all countries in the region, with the greater number of cases reported in Brazil, Peru and Chile<sup>8</sup>. As the curve of the coronavirus pandemic moves toward its highest peak in South America, there is an unprecedented need for mechanical equipment that provides oxygen to those who cannot breathe on their own and for health personnel trained to attend to patients arriving at the Intensive Care Units (ICU)<sup>9</sup>. However, Miguel Lago, director of the Instituto de Estudios para Políticas de Salud, IEPS, based in Río de Janeiro mentions that SA can be the region with the most number of victims due to Covid-19<sup>10</sup>.

Jesús Valverde, president of the Peruvian Society of Intensive Medicine says that for every 100,000 inhabitants, a country must have 10 ICU beds, or as minimum 6 beds, but several countries in the region were always below this number<sup>9</sup>. A comparative study published in the journal *Critical Care Clinics* in 2006, estimated the number of beds located in intensive care per 10,000 inhabitants around 2.9 in Argentina, 2.3 in Uruguay, 2.2 in Venezuela, 0.8 in Brazil, 0.3 for Chile, Colombia and Ecuador and 0.2 for Peru<sup>10</sup>. However, due to the fragmented nature of most healthcare systems in LAC, not all of these beds may be readily available to patients covered by public schemes. For example, in Brazil, only 40.6% of the total ICU beds are managed by the Sistema Único de Saúde (SUS), the health system financed with public funds. Similarly, in Ecuador and Paraguay, 53.2% and 41.4% of ICU beds, respectively, are present in the public sector of health systems<sup>8</sup>.

Although since 2006 to date the proportion may have risen in some countries ICU beds everything indicates that the number will be insufficient against a disease, that according to some estimates, could be contagious up to the 80% of the population and send to intensive care between 2% and 4% of those infected<sup>10</sup>.

## 1.2 Mathematical Model

A model is a representation of a process<sup>11</sup>. A mathematical model is a description of this process, in mathematical language, of an object that exists in a non-mathematical universe<sup>12</sup>. Mathematical

models are used in many fields of human activity, such as Mathematics, Economics, Physics, Chemistry, Biology, Psychology, Communication, Demography, Astronomy, Engineering, etc. Many practical problems require using mathematical models and sometimes situations are very different, but the approach and underlying philosophy are the same<sup>13</sup>.

Mathematical models have become a primary tool for learning, science development, prediction for decision-making and control regarding social and natural phenomena. Good decision making will help to obtain good results or in the opposite way it will produce great losses<sup>14</sup>. Usually, a mathematical model takes the form of a set of equations describing a number of variables, and we distinguish between continuous and discrete models. Some mathematicians have a procedure that they apply when they are building a model which is there is a phenomenon of interest that one wants to describe or, more importantly, explain<sup>11</sup>.

In general terms, in any mathematical model four phases can be determined in which the success or failure of these models is a reflection of the precision with which said mathematical model represents the initial object and not the exactitude with which mathematics analyzes the model<sup>12</sup>.

1. **Construction of the model.** Transformation of the non-mathematical object into mathematical language.
2. **Analysis of the model.** Study of the mathematical model.
3. **Model calibration.** A unique set of model parameters that provide a good description of the system behaviour
4. **Interpretation of mathematical analysis.** Application of the results of the mathematical study to non-mathematical initial object.

In resume, mathematical modeling is the link between something intangible (mathematics, physics, chemistry) and the real world. Is a process, A process, where you ask yourself a question, think a bit, and then you redefine the question, phrasing and placing it in precise mathematical. Once the question proposed becomes a mathematical question, as a scientific, a great field in mathematics opens up for you to carry out research that allows you to find an answer. Finally, is necessary that you revert the process, translating the mathematical solution back into comprehensible and with sense answer to the original question<sup>15</sup>.

### 1.3 General Classification of Mathematical Models

There are at least two different approaches to using models in research, each of which is chosen according to what the model is expected to be: deterministic and stochastic<sup>13</sup>. In a deterministic model, the factors involved in the study can be controlled by the process or phenomenon. Therefore it is possible to predict accurately the results obtained. In a stochastic model, it is not possible to control the factors involved in the study of the phenomenon. Consequently it does not produce simple unique results. Each of the results possible is generated with a probability function which assigns a probability to each of the trays; for example a model to predict the size of an epidemic in a population of  $N$  individuals. For the case deterministic provides a single value,  $C$ , while the stochastic model allows the possibility to obtain from zero to  $N$  individuals and is awarded a certain probability to each of these events. The difference is greater than it seems, since in a deterministic mathematical model in the context epidemiological; a single subject causes an epidemic generalized, while under a stochastic model there is a possibility that the epidemic will die out<sup>16</sup>.

It is important to mention that a model is defined by the relationships that it incorporates. These relationships are independent of the data to be analyze, since a model can be used to in different contexts and to different subjects. For example, in Epidemiology the application of mathematical models can be traced to the year 1760 when Daniel Bernoulli published a small investigation on the plague epidemic that was then sweeping Europe. In the last century, interest in the application of quantitative methods such as computational models to biology increased as a consequence of their success in physics and in particular in biophysics and biochemistry<sup>17</sup>.

### 1.4 Epidemic Mathematical Models

Mathematics and mathematical modeling are tools, languages and vanishing points, ways of seeing the world, which depend on the theoretical framework and empirical evidence that gives rise to them. Mathematics in epidemiology and in general in biological sciences constitute, moreover to being a tool, a way of thinking and structure predictions, descriptions, and explanations of processes. The language of mathematics allows to clarify and specify mechanisms, functions, and causal relationships between its components, which determine the evolution of a phenomenon

given<sup>17</sup>. The mathematical models of epidemiology are not isolated abstract constructions or they can be an end in themselves. There is always a theoretical framework of which each model, in particular, is justified since the hypotheses with those that are elaborated come from this and not from whose imagination builds it<sup>17</sup>.

Mathematical epidemiology was raised to a new level by the model of the spread of infectious diseases, published by Kermack and McKendrick in 1927 known as the SIR epidemic model<sup>18</sup>. Let  $S$  denote the number of individuals who are susceptible to the disease, that is, who are not infected at time  $t$ ,  $I$  denotes the number of infected individuals and  $R$  denotes the number of individuals who have been infected and then removed from the possibility of being infected again or of spreading infection<sup>19</sup>. Moreover, epidemiological models consist of systems of ODEs that describe the dynamics in each class. One of the simplest models involves the dynamics of susceptible, infectious, and recovered individuals (SIR)<sup>18</sup>. Within the SIR model, it is observed that its infectious behavior dynamics depend on the basic reproduction ratio  $R_0$ . Which is an epidemiological metric used to measure the transmissibility of infectious agents. However, this metric is an estimated value based on the early phase of the pandemic when the population has zero immunity. In addition,  $R_0$  is used to calculate the effective reproductive number  $R(t)$ , which aims to predict the probability of when the pandemic ends. For our case of study the parameter used is the growth rate ( $\lambda(t)$ ) which is also related to the transmissibility of infections. However, the estimation of  $R_0$  does not consider the parameters or health policies adopted by each country, since  $\lambda(t)$  is estimated from the total cases and  $R(t)$  is estimated from the assumption of  $R_0$ .

Although simple and powerful, mass action compartmental model SIR, does not capture the inherent heterogeneity of the underlying populations. Significant amount of research has been conducted to extend the model<sup>20</sup>. For example, the SEIR model describes the transmission dynamics of COVID-19 in China and forecasted the national and global spread of the disease, based on reported data from December 31, 2019 to January 28, 2020<sup>21</sup>. In this way, this model and others related to the COVID-19 spread can be helpful to do the research and modeling about the multi peaks found in South America.

## 1.5 Models

### 1.5.1 Exponential Model

One of the basic principles studied in mathematics is the observation of relationships between two connected quantities. A function is this connecting relationship, typically expressed in a formula that describes how one element from the domain is related to exactly one element located in the range<sup>22</sup>. The notion that knowledge grows exponentially seems to have first appeared in a short story by Sir Arthur Conan Doyle, “The Great Keinplatz Experiment,” which contains the statement, “Knowledge begets knowledge as money bears interest”<sup>23</sup>. According to the physicist Bartlett (1976), the mathematics of growth is the mathematics of the exponential function<sup>24</sup>. Exponential growth or decay is a consequence that follows whenever we have a function  $N$  which changes with time in such a way that the change  $\Delta N$  in  $N$  during a short time interval  $\Delta t$  is proportional to  $N$  and to  $\Delta t$ <sup>24</sup>. The equation we get is  $\frac{\Delta N}{\Delta t} = \lambda N$ . The time rate of change of the quantity is proportional to the quantity. The larger the value of  $N$  the faster it changes. The constant  $\lambda$  is fractional change ( $\Delta N/N$ ) in  $N$  per unit time  $\Delta t$ . Equation 1.1 is a differential equation which can be solved in several ways.

One way is  $N = N_0 e^{\lambda t}$  or in terms of function  $F(t) = N_0 e^{\lambda t}$  which are discussed in calculus terms in Appendix. Where  $N_0$  is the initial size or case, at time  $t$ ,  $e$  is the Euler’s number and  $\lambda$ , the continuous growth rate. The exponential function is not to be confused with the polynomial functions, such as  $x^2$ . One way to recognize the difference between the two functions is by the name of the function. Exponential functions are called so because the variable lies within the exponent of the function<sup>25</sup>. These functions are often recognized by the fact that their rate of growth is proportional to their value<sup>26</sup>.

### 1.5.2 Logistic Model

In an epidemic, cumulative incidence initially grows exponentially, but eventually slows and approaches a limit. This behavior is qualitatively similar to that of a logistic curve. Thus, a logistic model may allow us to use longer sequences of data from the beginning of an epidemic, by accounting for the epidemic slowing as it proceeds<sup>27</sup>. The logistic function finds applications in a range of fields, including biology (especially ecology), biomathematics, chemistry, demogra-



phy, economics, geoscience, mathematical psychology, probability, sociology, political science, linguistics, statistics, and artificial neural networks<sup>28</sup>.

The logistic function was invented in the 19th century for the description of the growth of populations and the course of autocatalytic chemical reactions<sup>29</sup>. The Logistic model originated from the modeling of population growth in ecology. As an improvement on the Malthus population model, in 1838, Pierre Franois Verhulst published the logistic equation<sup>3</sup>. The way that the logistic growth occurs is a little bit easy to explain because is characterized by an increasing curve in the beginning period, but in a time the curve starts decreasing, as you get closer to a max.

### Discrete logistic function

In terms of population growth, this fact ranges from exponential growth, which will eventually decline due to resource factors or competition. With this in mind, it is suggested that the assumption that growth will be equal for the rest of the days should be changed regardless of the current population size. That is, the estimated number for the next few days will depend on the size of the population but in various forms.

A generalization of the logistic function or logistic curve is the hyperbolastic function of type I, a common S-shaped curve (sigmoid curve) with equation;

$$f(t) = N / (1 + e^{-\lambda(t-t_0)})$$

where  $t_0$  is the valune of the sigmoid's midpoint,  $N$  is the curve's maximum value and  $\lambda$  is the logistic growth rate.

### Applications

In ecology, a typical application of the logistic equation is a common model of population growth, where the rate of reproduction is proportional to both the existing population and the amount of available resources, all else being equal<sup>30</sup>. In medicine, a logistic function, or related functions (e.g. the Gompertz function) are usually used in a descriptive or phenomenological manner because they fit well not only to the early exponential rise, but to the eventual levelling off of the pandemic as the population develops a herd immunity<sup>31</sup>. This is in contrast to actual models of pandemics which attempt to formulate a description based on the dynamics of the pandemic

(e.g. contact rates, incubation times, social distancing, etc.). Some simple models have been developed, however, which yield a logistic solution<sup>32,33</sup>.

### **Logistic differential equation**

In fact, another way to take into the logistic equation is in its differential form which depends on the number of cases or on time. The use of this ODE allow us explain the irregularities in the number of cases per day, and the appearance of several peaks of incidence and their duration, which are not predicted by the deterministic or generalized logistic equation. Our goal is not to solve the equation, but to investigate the relationship between the number of cases per day and the total number of the infected people, which is expressed by the algebraic curve<sup>34</sup>.

## **1.6 General and Specific Objectives**

### **1.6.1 General Objective**

- Develop a suitable modeling of the spread of COVID-19 based on the logistic model.

### **1.6.2 Specific Objectives**

- Propose a simple mathematical approach using the exponential model as well as the discrete and differential logistic equations.
- Study the effect of COVID-19 outbreak in South America
- Apply the logistic model to analyze the spread of COVID-19 in Ecuador

# Chapter 2

## Methodology

### 2.1 Theoretical Approach

We shall start with the following simple model. Exponential functions can be used to model population growth, interest rates, radioactive decay, and the amount of medicine in the bloodstream. Since our interest is to model the increasing number of infected people during the pandemic, the following equation is proposed:

$$\frac{dN(t)}{dt} = \lambda_{exp} N(t) \quad (2.1)$$

Exponential growth rate  $\lambda_{exp}$  is taken as the positive growth rate of the epidemic ( $day^{-1}$ ) and  $N(t)$  represents the total number of infected people which depends on time. Solving step-by-step Equation 2.1:

$$\int \frac{dN(t)}{N(t)} = \int \lambda_{exp} dt \quad (2.2)$$

$$\ln N(t) = \lambda_{exp} t + c \quad (2.3)$$

$$N(t) = e^{\lambda_{exp} t + c} \quad (2.4)$$

$$N(t) = e^{\lambda_{exp} t} e^c \quad (2.5)$$

one can finally get

$$N(t) = N_0 e^{\lambda_{exp} t} \quad (2.6)$$

where  $N_0$  is the initial number of the infected people and  $t$  is the time since the first reported cases (day). Equation 2.6 can only be used at the early stage of the pandemic if the number of confirmed COVID-19 cases as a function of time is described by a clear exponential equation, which is illustrated in Figure 1a. In addition, during this initial period, we don't observe the depletion of possible infection candidates.

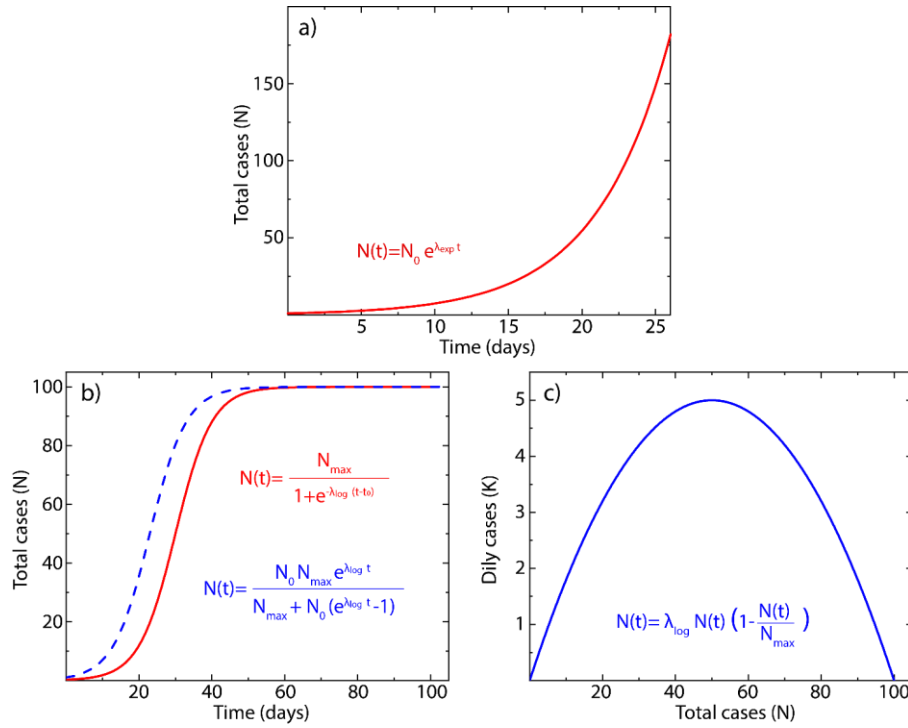


Figure 2.1: Illustration cases: a) the total number of cases as a function of time using the exponential model, b) the total number of cases as a function of time. The discrete logistic equations (red line) are compared to the solution of the ODE logistic model (dashed blue line), and c) the daily cases as a function of total cases. Parameters used:  $N_{max} = 100$ ,  $N_0 = 1$ ,  $\lambda_{log} = 0.2(day - 1)$ ,  $\lambda_{exp} = 0.2(day - 1)$ ,  $t_0 = 30 (day)$ .

In fact, this model can be applied as long as there are no changes in the curve. On the other hand, if restrictions are implemented to flatten the curve (such as social distancing, case testing, hand washing, and quarantine), the dataset must be fitted using the discrete logistic equation:

$$N(t) = \frac{N_{max}}{1 + e^{-\lambda_{log} (t-t_0)}} \quad (2.7)$$

which also allows estimating the maximum number of infected people ( $N_{max}$ ), the logistic growth rate ( $\lambda_{log}$ ,  $day^{-1}$ ), and the point of maximum growth ( $t_0$ ). Logistic functions were first studied in the context of population growth, as early exponential models failed after a significant amount of time had passed. By adding the correcting factor  $\frac{\lambda_{log} N(t)}{N_{max}}$  in Equation 2.1 and taking  $\lambda_{exp} = \lambda_{log}$ , the resulting equation is:

$$\frac{dN(t)}{dt} = \lambda_{log} N(t) \left( 1 - \frac{N(t)}{N_{max}} \right) \quad (2.8)$$

Equation 2.8 is the logistic equation written in the form of ODE.  $N(t)$  denotes the total number of confirmed cases as a function of time, and the respective solution is easily found as follow (For extend solution check Appendix A.2):

$$N(t) = \frac{N_0 N_{max} e^{\lambda_{log} t}}{N_{max} + N_0 (e^{\lambda_{log} t} - 1)} \quad (2.9)$$

Note that Equation 2.7 and Equation 2.9 can be used to estimate the  $\lambda_{log}$  and  $N_{max}$  parameters from the curve of the total case (Figure 2.1b), but the  $t_0$  parameter cannot be estimated from Equation 2.9. Now, using the difference logistic equation as follow:

$$\frac{dN(t)}{dt} = N_{i+1} - N_i = K(t) \quad (2.10)$$

where  $K(t)$  is the daily reported cases, the following expression is found:

$$K(t) = \lambda_{log} N(t) \left( 1 - \frac{N(t)}{N_{max}} \right) \quad (2.11)$$

which shows a simple relation between  $K(t)$  (daily cases) and  $N(t)$  (total cases) (Figure 2.1c). Interestingly enough, Equation 2.11 can also be used to estimate the  $\lambda_{log}$  and  $N_{max}$  parameters. Additionally, the logistic growth rate can be found from Equation 2.11, as follow:

$$\lambda_{log} = \frac{K(t)}{N(t) \left(1 - \frac{N(t)}{N_{max}}\right)} \quad (2.12)$$

in fact, Equation 2.12 can be used to study the growth rate variability as a function of total confirmed cases ( $N(t)$ ) or time ( $t$ ). To simplify the notation,  $\lambda_{exp}$  and  $\lambda_{log}$  have been taken as  $\lambda$  through the text, Figures, and Tables.

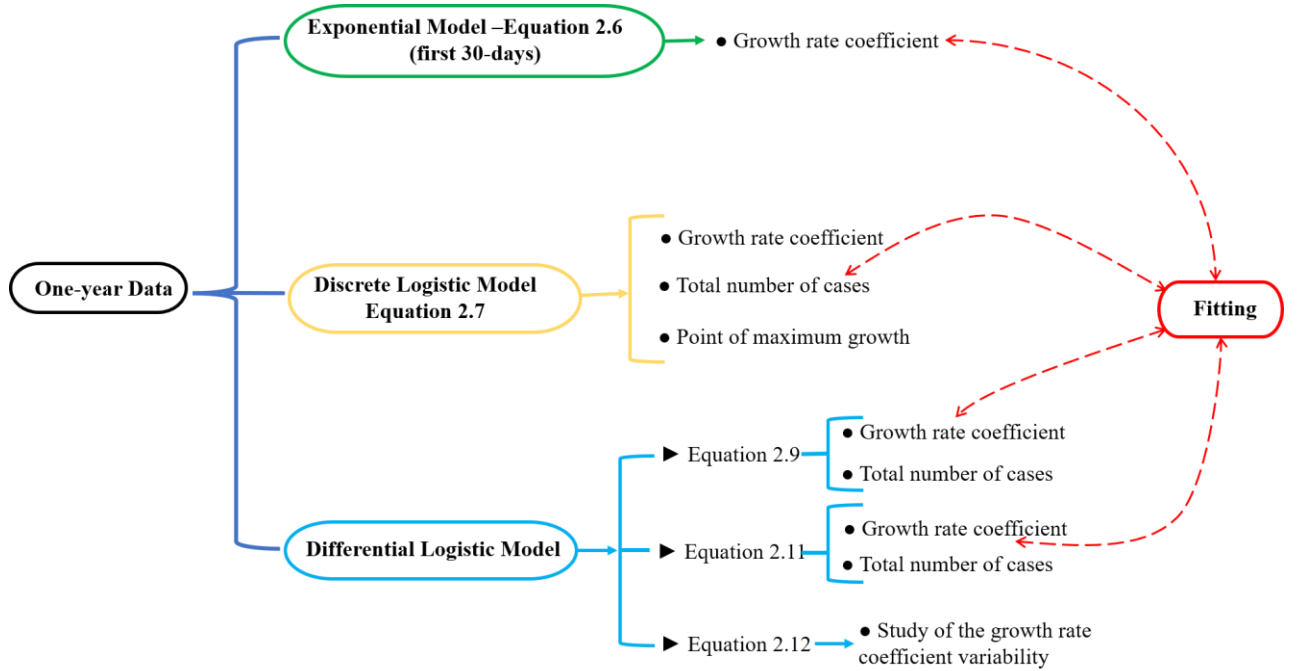


Figure 2.2: Schematic representation of the research methodology

Figure 2.2 shows the research methodology of the present work. In the exponential model, the data of the first 30 days were used, where an exponential increase in the number of infected people was observed. In the other models, the complete data set was analyzed with a seven-day moving average (359 data points). As observed, the exponential model gives access to the exponential growth rate at the beginning of the pandemic. On the other hand, the discrete logistic model helps correct the error of the exponential model for long periods, also allowing to estimate the logistic growth rate, the total number of infected people, and the inflection point of the curve. Most importantly, the differential logistic model can give a more complete picture of the COVID-19

pandemic because Equations 2.9 and 2.11 can be used to estimate the growth rate coefficient and the total number of infected people as a function of time or total cases, respectively. Finally, Equation 2.12 allows to study the growth rate coefficient variability, evidencing the random phenomenon of data as a consequence of the collecting/reporting data or social behavior.

## 2.2 Data Collection and Analysis

One-year data of COVID-19 outbreak in South America, from March 1, 2020, to February 28, 2021, was taken (and compared) from:

- Our World in Data: Coronavirus Pandemic (COVID-19)  
(<https://ourworldindata.org/coronavirus>)
- Worldometer  
(<https://www.worldometers.info/coronavirus>)
- Johns Hopkins Coronavirus Resource Center  
(<https://coronavirus.jhu.edu>)

From these sources, we have extracted the following values:

- Pandemic start cases and days.
- Accumulated cases in a year (The total number of coronavirus cases).
- Daily new reported cases in a year.

Some discrepancies were found between the given data sources, particularly, at the beginning of the pandemic. As mentioned, this can be attributed to the lack of qualified laboratories to carry out tests and large inexperience in collecting and reporting daily data. This problem is present in every country studied in this work. For this reason, the study is carried out using a seven-point (7p) moving average, which means that it takes the last 7 days, adds them up, and divides them by 7. This smooth-out resulting curves cancel out peaks and valleys in data collection. Hence, the seven-point moving average list was used in the exponential and logistic models. All parameters considered in our models are:



- The total number of infected people as a function of time
- The initial number of the infected people  $t=0$  days
- The maximum number of infected people
- The point of maximum growth

# Chapter 3

## Results & Discussion

In this chapter, we develop the proposed models for the modeling of the spread of COVID-19 and the goal is to explain the main characteristics of each one and what we obtained from them. We start with exponential model with a minimum range of values because it is applied at the early days (30) of the pandemic. We also explore the values for  $\lambda$ . We compare the graphs built from the data and results to find a relationship between the model and the first thirty days expectations. Next, in a wide case we analyze the total cumulative cases with the discrete and differential logistic model. In the same parameters as growth rate and total estimated cases.

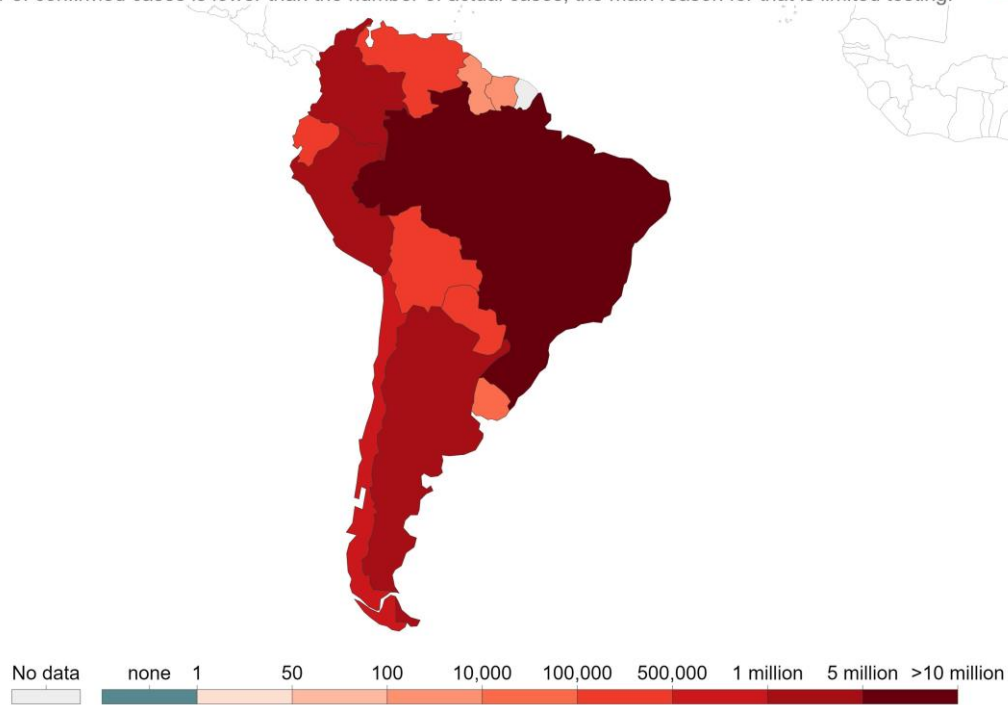
### 3.1 South America daily and cumulative cases

South America is currently an epicenter of the COVID-19 pandemic. The impact of the pandemic in these countries is quite diverse, having affected mainly countries like Brazil and Peru. The spread of COVID-19 in South America is shown in Figure 3.1.

### Cumulative confirmed COVID-19 cases

The number of confirmed cases is lower than the number of actual cases; the main reason for that is limited testing.

Our World  
in Data



Source: Johns Hopkins University CSSE COVID-19 Data

CC BY

Figure 3.1: Cumulative cases map of confirmed COVID-19 cases until Feb 28, 2021. Source: COVID-19 Data Repository by the Center for Systems Science and Engineering (CSSE) at Johns Hopkins University ( <https://github.com/CSSEGISandData/COVID-19>).

Reported data are shown in the black color and 7-point moving average data in the red color. As mentioned earlier, the 7-point moving average dataset will be used throughout this work to avoid the peaks and valleys due to inadequate daily data collection and reporting. Although South America had about two months to prepare for the pandemic from the beginning in China, the health system and communication channels could not fulfill their purpose against the COVID-19 disease as observed in Figure 3.2, Figure 3.3, Figure 3.4 and Figure 3.5 and discussed below.

The reported data obtained from our world in data and plotted in Wolfram Mathematics as mentioned have a moving average correction of 7 points. The main cases for which this correction occurs is first because it seeks to remove the points in which daily cases with values of zero are reported and second is for negative values, as is the case of Ecuador and Guyana. The explanation for this situation in which the values are negative, is due to the mishandling of the data reported by these governments. The moment they reduce a figure in the confirmed cases is to show that they incorrectly carried the use of the detection tests, the complication exists at the moment that the results have time of delay in the delivery of them, false positives and are sometimes attributed as cause of death to people who were given the test but it was too late.

The figures are explained in such a way that, all those that are on the left side correspond to the number of daily cases of infected people, and all those images that are on the right side correspond to the number of accumulated cases of infected people. In addition, all the data of the Y axis of the figures are about 10000 or  $10^4$ , to handle numbers that are explained in a better way.

In the first instance we can analyze the daily cases (figures on the left side), which are based on the time in a window of 365 days from their first confirmed case. It can be seen, the data reported are characterized by intense peaks in their first days, whatever country do you want to analyze. Each country at least consists of two peaks between its first 100-150 days and a second after 150-200 days more. Looking at it by date the increases exist once the quarantines ended. As soon as WHO declares on 11 March, 2020 its deep concern about the bitter levels of spread, it enters into evaluation and evaluates that COVID-19 should be characterized as a pandemic. For this reason, all South American countries, enter a total confinement, for around 60 days or more (>2 months). Once this restriction ends the general population begins to take fewer precautions and the excessive increase occurs. That as we mentioned will be our first peak of infections in the figures of daily cases for each country.

The second peak is explained by the reason that on December 31, 2020, (WHO) declares in a press release the validation of a vaccine against COVID-19. Which is allowed worldwide in cases of emergence and supply to the general population. This unbinds a series of events in which, the misinformation prevail. Believing that such a vaccine "cured" or toward assuming people who were vaccinated that they would not have that disease only by the fact of getting vaccinated and again begin the careless ones that trigger the second peak or second outbreak in the population.

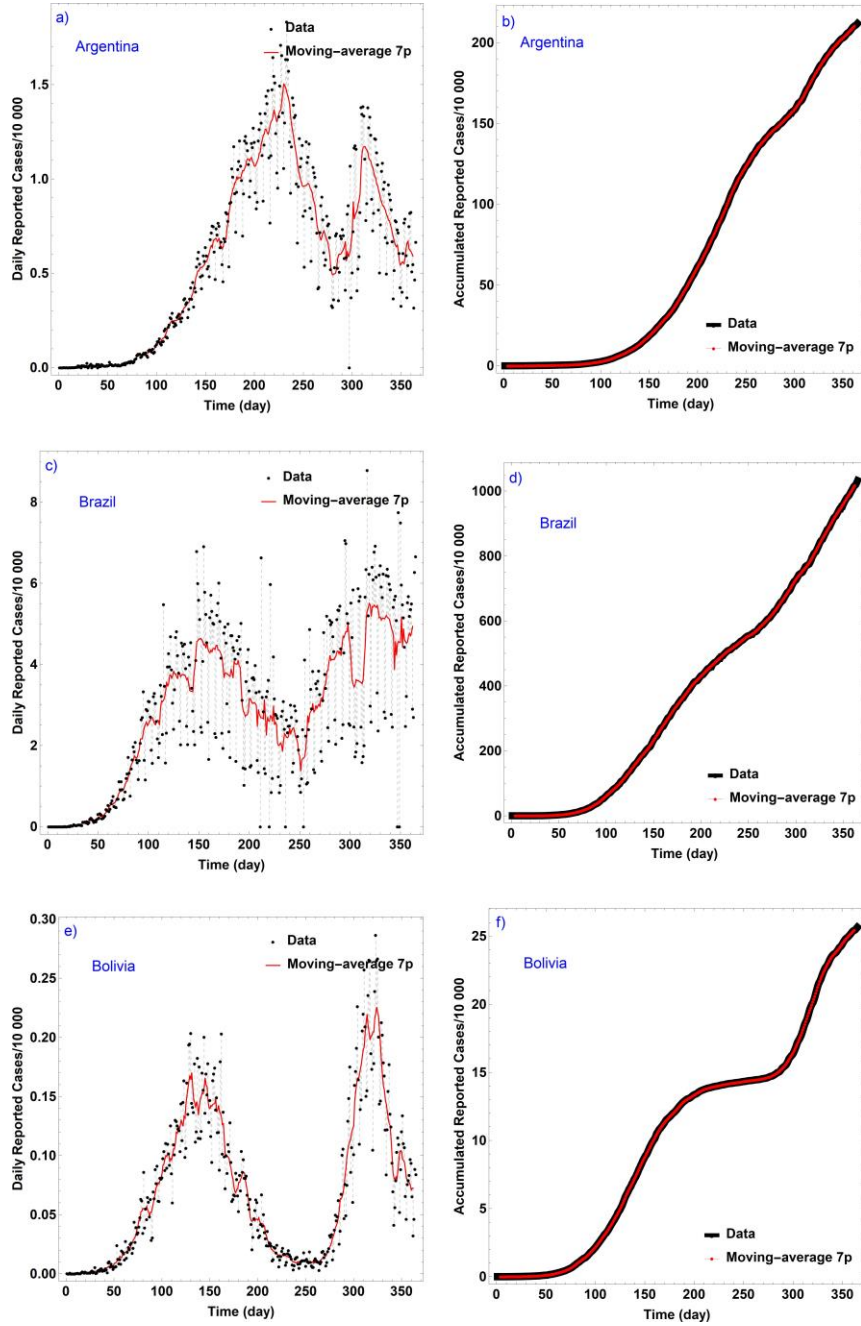


Figure 3.2: a) the number of infected people per day in Argentina, b) the total number of confirmed cases reported during 365 days in Argentina, c) the number of infected people per day in Brazil, d) the total number of confirmed cases reported during 365 days in Brazil, e) the number of infected people per day in Bolivia, f) the total number of confirmed cases reported during 365 days in Bolivia. Data are represented in black and the 7-point moving average data in red.

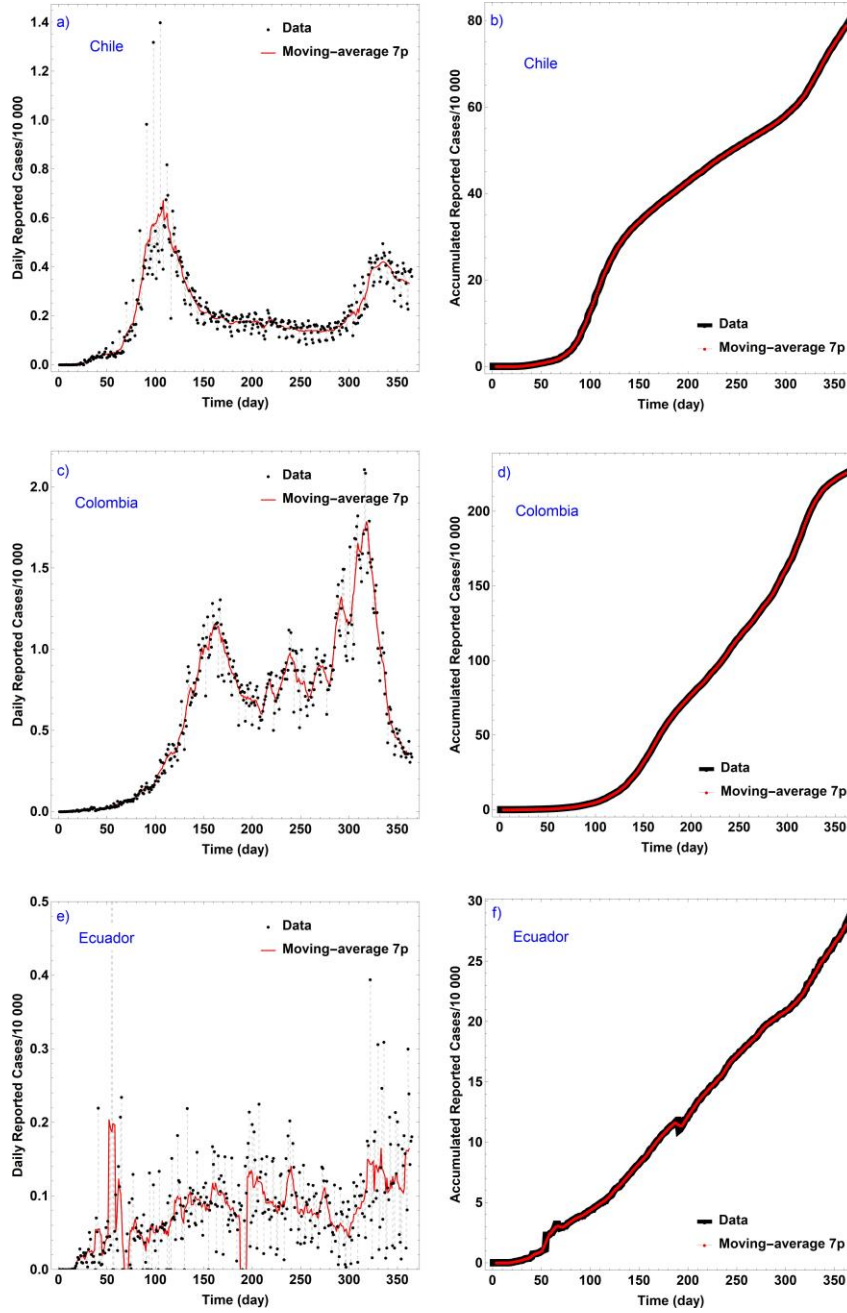


Figure 3.3: a) the number of infected people per day in Chile, b) the total number of confirmed cases reported during 365 days in Chile, c) the number of infected people per day in Colombia, d) the total number of confirmed cases reported during 365 days in Colombia, e) the number of infected people per day in Ecuador, f) the total number of confirmed cases reported during 365 days in Ecuador. Data are represented in black and the 7-point moving average data in red.

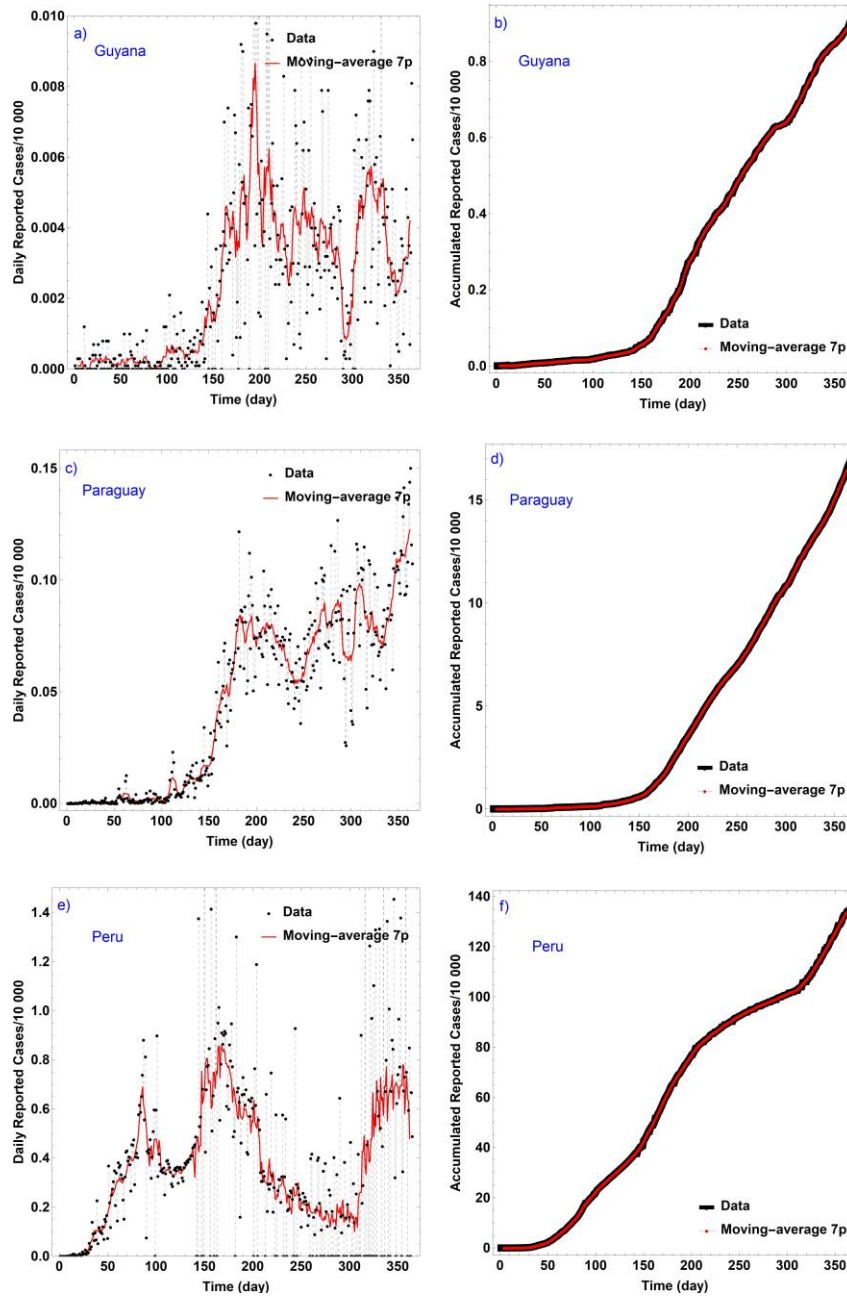


Figure 3.4: a) the number of infected people per day in Guyana, b) the total number of confirmed cases reported during 365 days in Guyana, c) the number of infected people per day in Paraguay, d) the total number of confirmed cases reported during 365 days in Paraguay, e) the number of infected people per day in Peru, f) the total number of confirmed cases reported during 365 days in Peru. Data are represented in black and the 7-point moving average data in red.

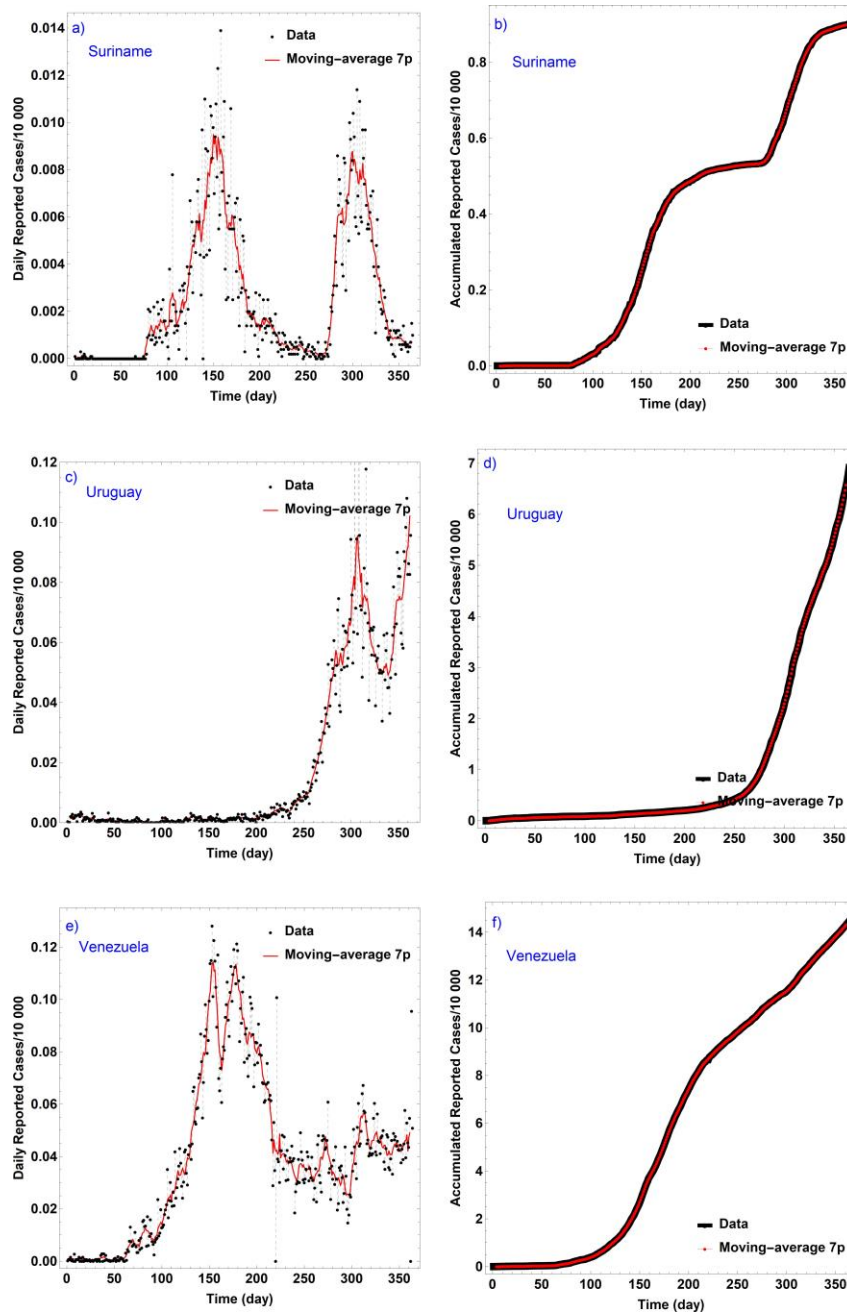


Figure 3.5: a) the number of infected people per day in Suriname, b) the total number of confirmed cases reported during 365 days in Suriname, c) the number of infected people per day in Uruguay, d) the total number of confirmed cases reported during 365 days in Uruguay, e) the number of infected people per day in Venezuela, f) the total number of confirmed cases reported during 365 days in Venezuela. Data are represented in black and the 7-point moving average data in red.



## 3.2 Exponential Model

Unable to have an effective solution at a pandemic global level, modeling is an excellent option to explore the effect of COVID-19 and predict its consequences. In this context, as a first idea we had to predict it with an exponential modeling or exponential growth rate. From this, the cumulative data of infected people were adjusted with a 7-point moving average for the exponential model to be applied (Equation 2.6). This model is used to describe the early trajectory of this infection. This equation proposes to look for an exponential growth rate using the data accumulated from the day the first case was presented  $N_0$ .  $\lambda$  represents the growth rate that describes how quickly the infection spreads when restrictions (physical distancing, mask wearing, environmental cleaning, home disinfection, isolation, etc.) are not implemented to counter the spread of COVID-19.

It is important to mention that the exponential model, being a basic type modeling should only be adjusted for the first 30 days as shown on the X-axes of the Fig. 3.6 and Fig. 3.7. This due to the model stops working beyond this amount, because after these days certain limitations could be taken to avoid its rapid contagion, as predicted by the model.

Once the results of the modeling of each country were obtained, a table of the parameters found was made, for a better visualization and analysis of the same, that is Table 3.1. The most important point to note in all the cases that the modeling was performed, is corresponding to its  $\lambda$ , i.e. its growth rate, this value ranges from 0,073 – 0,258  $day^{-1}$ . These high values of  $\lambda$  are attributed to the fact that at the beginning of the pandemic the controls and restrictions were poor and health policies were not implemented quickly and in time to avoid the rapid spread of the infection. Nevertheless, not all the countries follow the same rhythm at the moment of bringing the contagions of COVID-19 and of taking these measures. For example, Brazil that because it was the first country to report contagions and with the largest population in South America, its number of infected was going to grow faster compared to other countries in the continent.

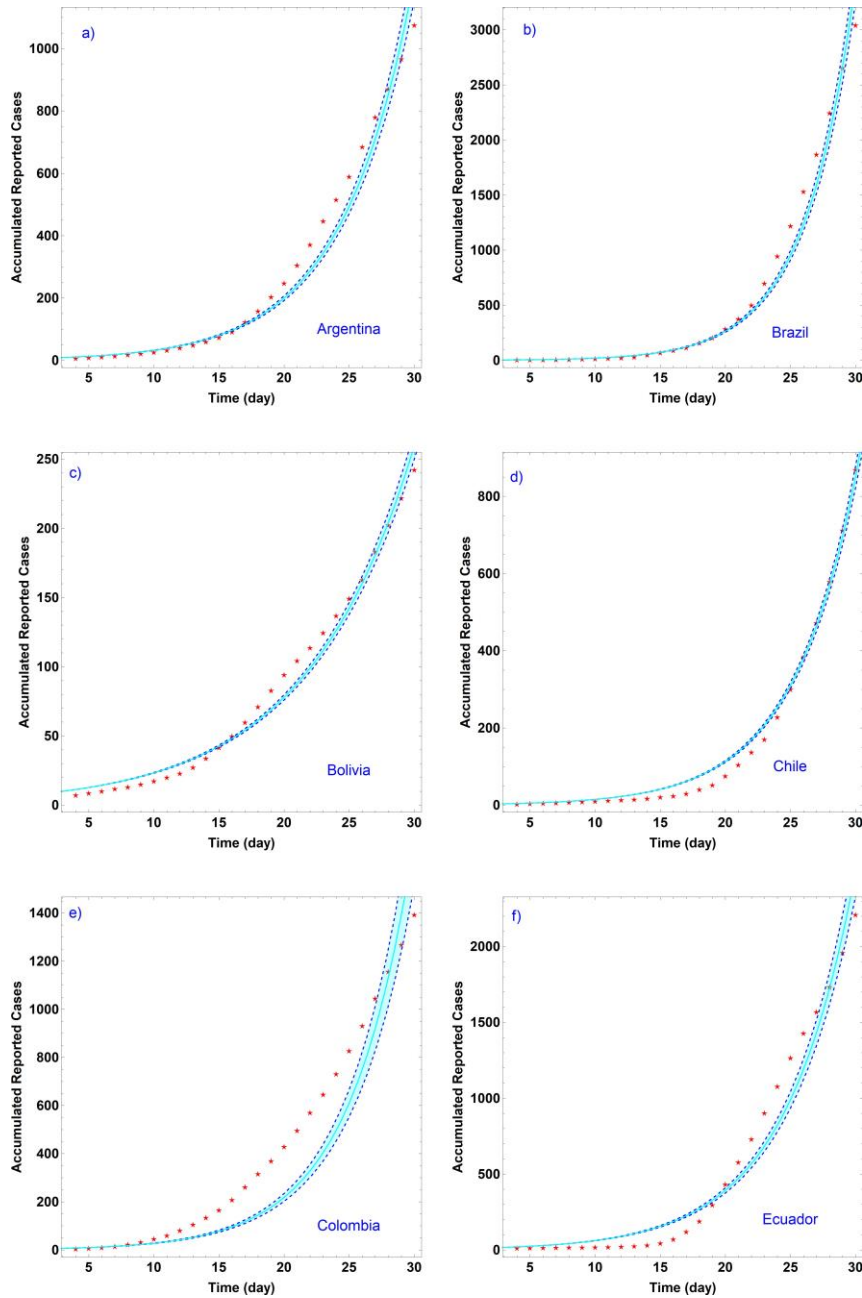


Figure 3.6: Exponential model using 30 days with 95% of confidence level. a) Argentina, b) Brazil, c) Bolivia, d) Chile, e) Colombia, f) Ecuador. Total cases data as function of time (red points) and fitting model with bands (blue lines).

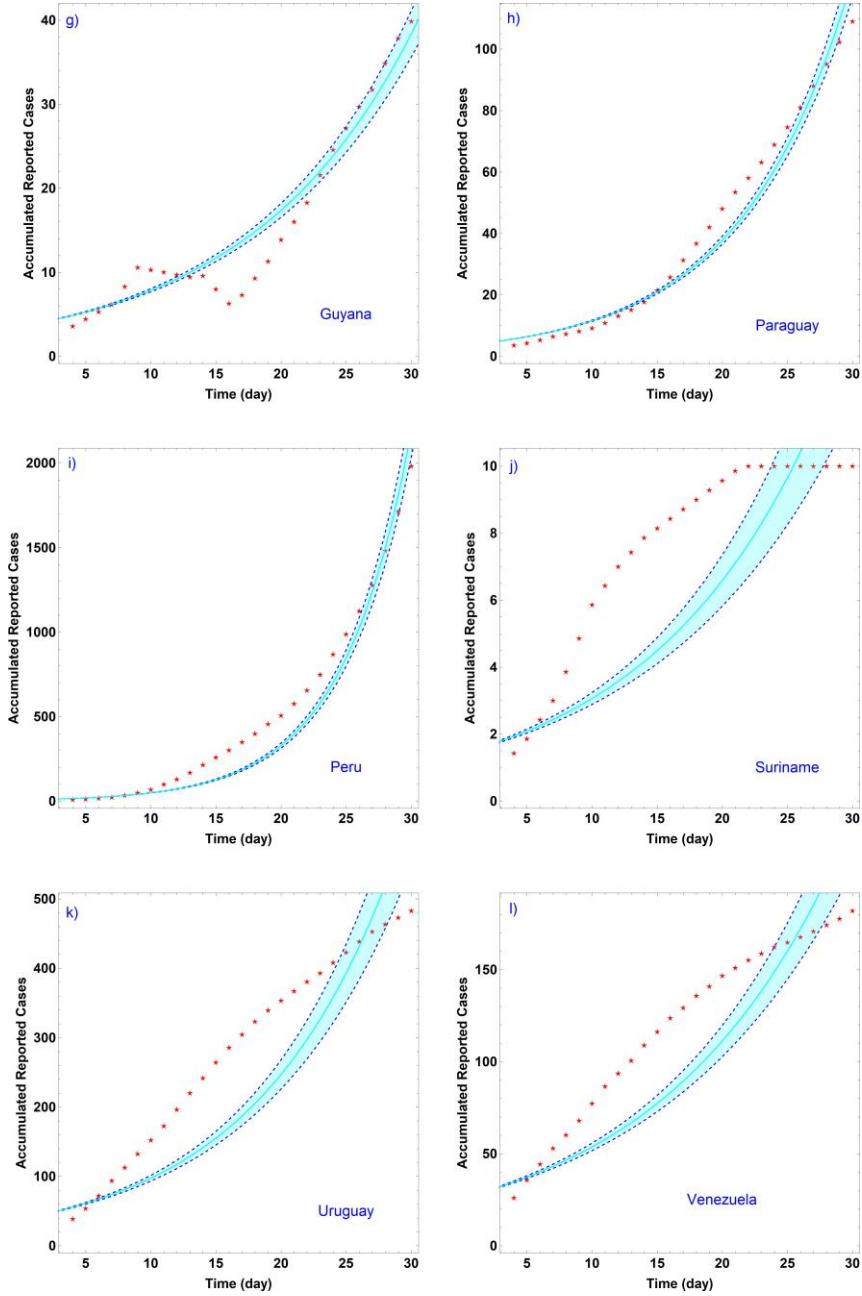


Figure 3.7: Exponential model using 30 days with 95% of confidence level. g) Guyana, h) Paraguay, i) Peru, j) Suriname, k)Uruguay, l) Venezuela. Total cases data as function of time (red points) and fitting model with bands (blue lines)

Table 3.1 has three important parameters that need to be analyzed. First, the growth rate ( $\lambda$ ) obtained through the application of the model with its respective standard error. The second parameter is the curve fit ( $R^2$ ), which shows how well the curve obtained from the modeling fits to the real data, whether daily or accumulated. According to BBC news, Peru was the second most contagious country at the start of the pandemic, but it was one of the first to react with strict containment and identification measures for the virus in its population. But part of the explanation for this is the high number of tests that were carried out by the Peruvian authorities<sup>35</sup>. But on the contrary, we have the case of Ecuador, which according to the data simulated that in a single day the number of infected people doubled, This is because the country obtained the results of thousands of COVID-19 tests that were delayed, as reported by the Minister of Health, Juan Carlos Zevallos, in a press conference<sup>36</sup>.

Exponential Model Results			
Country Name	Growth rate ( $\lambda$ )	Standard Error( $\pm\lambda$ )	$R^2$
Argentina	0.182	0.001	0.984
Brazil	0.256	0.001	0.984
Bolivia	0.120	0.001	0.993
Chile	0.202	0.001	0.994
Colombia	0.206	0.002	0.939
Ecuador	0.185	0.001	0.978
Guyana	0.079	0.001	0.978
Paraguay	0.118	0.001	0.987
Peru	0.190	0.001	0.976
Suriname	0.076	0.003	0.898
Uruguay	0.093	0.002	0.937
Venezuela	0.073	0.002	0.952

Table 3.1: Estimated growth rate parameter ( $\lambda$ ) from the exponential model for the different countries under study and curve fit ( $R^2$ ) for each one.

### 3.3 Discrete Logistic Model

In the area of growth dynamics, the use of the discrete logistic model is mentioned continuously. The reason why this model is called as discrete model is because an analysis of the growth rate is not performed continuously. It means, that the growth rate is counted one by one, for example, from the day 1 at 8 a.m. to the day two at 8 a.m. and so on for the following days. In other words, this model is called discrete due to instead of taking all possible values over time (8:30 a.m, 9:00 a.m, and so on) continuously from 0 to 365 days, we are taking it one by one (day 1, day 2, day 3 and so on). Unlike the differential logistic model where we clearly are in continuous domain. Moreover, this model is a representation of a general population growth function and can be closely approximated to birth/death dynamics and processes with similar characteristics<sup>37</sup>. On a practical level, due to complex behaviors in ecology, the equation can be transformed into its deterministic or discrete form<sup>38</sup>.

In terms of pandemic, the logistic equation has been taken into account because many of its parameters can be obtained easily, only using the accumulated data either several months or annually. For our case the accumulated data are taken annually and in the same way with a 7-point moving average and the discrete equation of the logistic model is applied (Equation 2.7).

The results of this equation are the estimated value of the total cases accumulated within what could be given in the year  $N_{max}$ . The growth rate  $\lambda_{log}$  for this model will be explained later. In addition, the midpoint of the spread  $t_0$ , that is, on which day the value of  $N_{max}$  would be almost reached. These results are tabulated in Tables 3.6 and 3.3. And the graphs concerning the modeling of the logistic equation in the Figures 3.8 and 3.9.

The discrete logistic model, shows an excellent agreement between the reported data and the curve corresponding to the model in all the cases analyzed. This is proven by Table 3.3, in which the  $R^2$  has an index greater than 0.97, which is more than enough accepted for analysis. To this is added the check by the use of the interval of confidence (intermittent blue line), in which are some data below or above these. Although in all cases the fitting curve, has a  $R^2 > 0.8$ , the values estimated by this equation for the year are not always correct. In the cases of Argentina, Bolivia, Chile, Guyana, Peru and Venezuela, values are underestimated. This means that in the year evaluated the actual total cases far exceeded the estimated total cases. About 50% of the countries evaluated have values close to those estimated. It is important to mention that this model should be used carefully at the time of predicting values for the year and even for dates

after the year.

Discrete Logistic Model Results					
Country Name	Total real cases	Estimated cases $N_{max}$	Standard error( $\pm N_{max}$ )	Estimated time $t_0$	Standard Error( $\pm t_0$ )
Argentina	2 118 676	2 099 630	13 210	239.491	0.822
Brazil	10 390 461	10 891 700	209 280	242.930	3.221
Bolivia	256 462	244 260	6 800	211.378	5.097
Chile	803 009	695 250	9 870	175.399	2.615
Colombia	2 269 582	2 701 910	44 630	266.846	2.440
Ecuador	286 155	304 050	3 960	230.566	2.433
Guyana	8 993	8 910	80	244.242	1.169
Paraguay	168 043	179 680	2 280	274.712	1.586
Peru	1 349 847	1 176 770	9 270	177.087	1.377
Suriname	9 022	9 090	210	215.666	3.958
Uruguay	69 074	77 420	1 040	322.572	0.876
Venezuela	144 786	130 870	880	198.193	0.992

Table 3.2: Estimated cases ( $N_{max}$ ) and midpoints ( $t_0$ ) parameters with their standard errors from the discrete logistic model for the different countries under study.

It is important to analyze, the midpoints of each country ( $t_0$ ) in Table 3.6 since this value shows us the approximate dates and the approximate time in which each country suffered a flattening of cases, that is to say it was possible to decrease at some point the spread. Although not all cases are similar, for reasons such as population numbers, availability of hospitals and health measures. Almost all of South America, has a flattening within 200 days after the start of the pandemic. As mentioned this is due to the sanitary measures taken in each country and that in many cases were not given immediately, but took approximately 6 to 7 months.

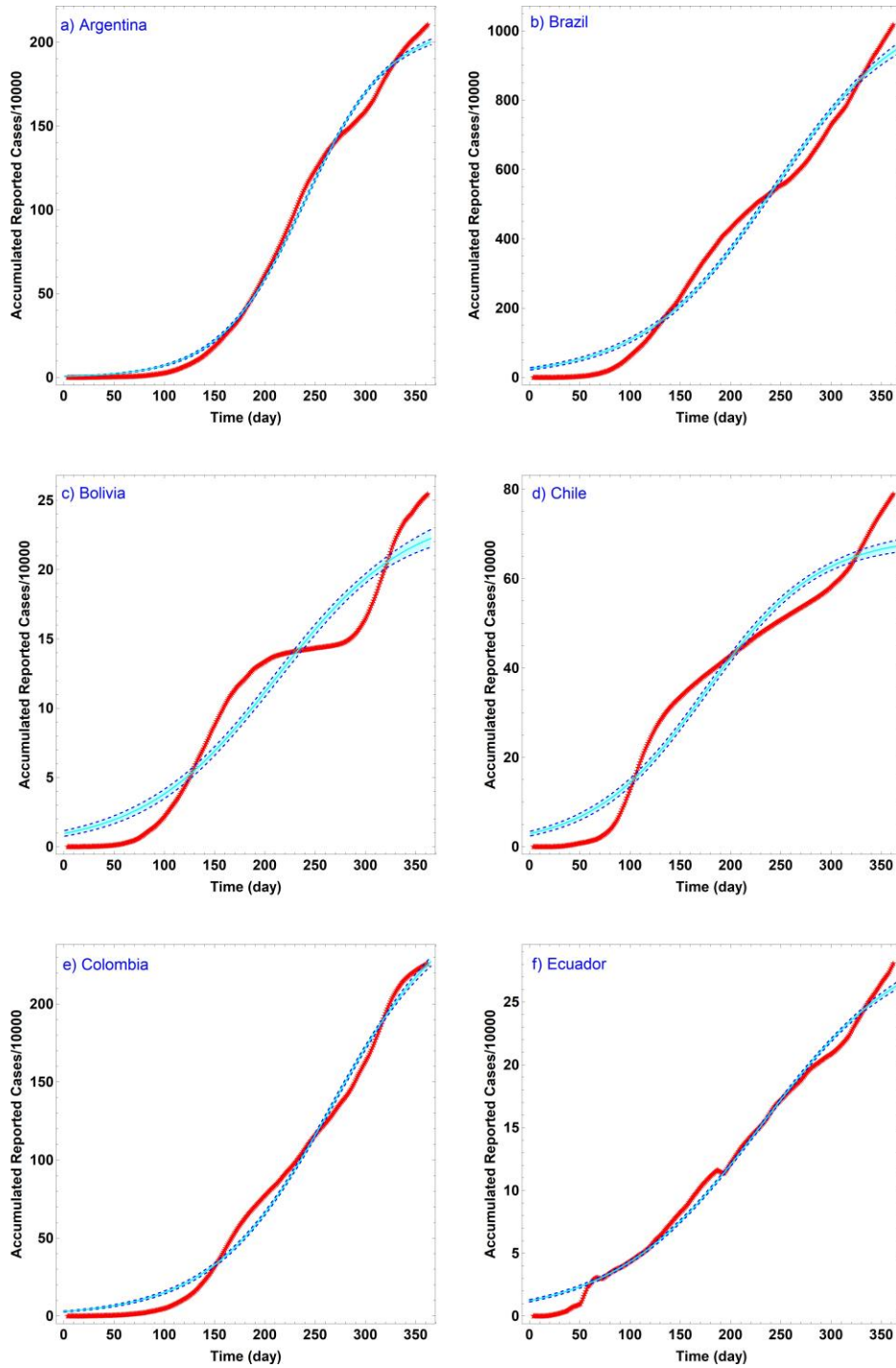


Figure 3.8: Discrete logistic model using moving average data with 95% of confidence level. a) Argentina, b) Brazil, c) Bolivia, d) Chile, e) Colombia, f) Ecuador. Total cases data as function of time (red points) and fitting model with bands (blue lines).

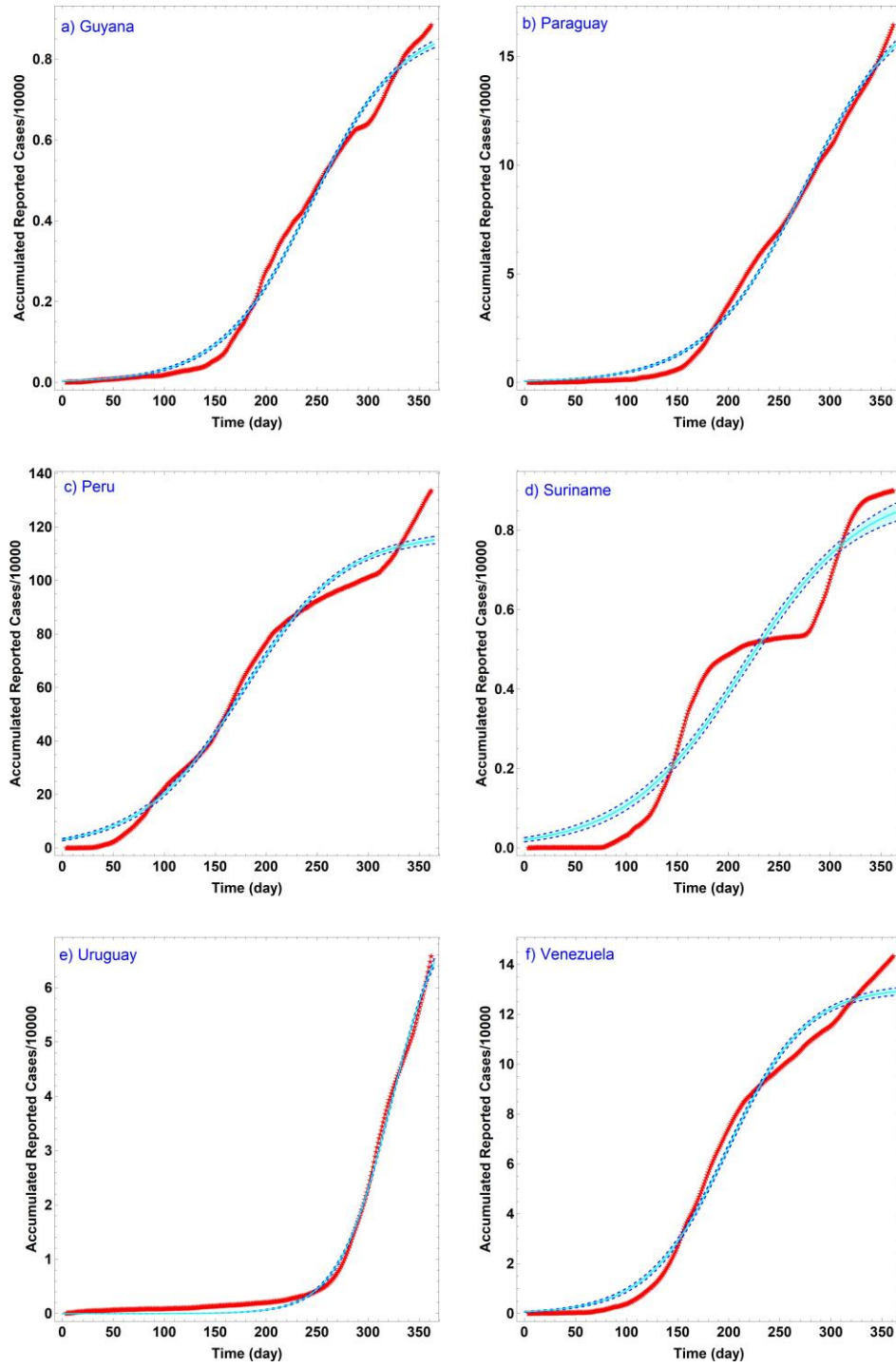


Figure 3.9: Discrete logistic model using data moving average with 95% of confidence level. a) Guyana, b) Paraguay, c) Peru, d) Suriname, e) Uruguay, f) Venezuela. Total cases data as function of time (red points) and fitting model with bands (blue lines).



Discrete Logistic Model Results			
Country Name	Growth rate ( $\lambda$ )	Standard Error( $\pm\lambda$ )	$R^2$
Argentina	0.024	0.000	0.998
Brazil	0.015	0.000	0.992
Bolivia	0.015	0.001	0.976
Chile	0.018	0.001	0.986
Colombia	0.017	0.000	0.995
Ecuador	0.014	0.000	0.997
Guyana	0.023	0.000	0.996
Paraguay	0.021	0.000	0.997
Peru	0.020	0.000	0.994
Suriname	0.017	0.001	0.977
Uruguay	0.038	0.001	0.997
Venezuela	0.026	0.001	0.995

Table 3.3: Estimated growth rate parameter ( $\lambda$ ) from the discrete logistic model for the different countries under study and curve fit  $R^2$  for each one.

### 3.4 Differential Logistic Model

In a previous section we discussed a model of population growth in which the growth rate is proportional to the size of the population (exponential model). In the resulting model the population grows rapidly and collapses as soon as possible. This is because the model is unrealistic and imposes environments limitations to population growth. Here it is proposed a more accurate model which postulates that the relative growth rate decreases when the total number of infected people approaches the carrying capacity  $N_{max}$  of the environment. The corresponding equation is the so called differential logistic equation (Equation 2.11).

Figure 3.10 and 3.11 show the behavior of the model curve applied to the countries studied. In the first instance, we can find a discrepancy in all cases in terms of following the trend of real data and fitting curve. What translates as the differential logistic model should be applied with extreme caution when estimating values for a long period of time. This is verified with the use of Table 3.4, in which the estimated values for the  $N_{max}$  are present. Although with the use of Equation 2.9, it is not evaluated when a flattening in the curve will occur ( $t_0$ ). We can clearly distinguish when countries collapse. If we compare Figures 3.10 and 3.11 with the data in Table 3.2, we can find that in these figures flattening occurs on approximately the same days where flattening was predicted. Once we find this phenomenon present in the figures of this model, we also visually realize that the estimated values of the total cases accumulated with the use of this model is extremely lower. That is, in all cases the estimate is very inefficient and inaccurate. Since once the flattening is found, the model assumes that the growth should stop, something that did not happen in real life. This flaw can be compared to the values of Equation 2.7. In many cases this failure exceeds by almost half of the values found.

For example for Ecuador  $N_{max} = 203199$  (differential logistic model) and  $N_{max} = 304050$  (discrete logistic model), only in this case we realize that it exceeds by at least 100000 cases between models. In a more extreme case we have Brazil, being one of the most affected countries in South America. The estimated value with the differential logistic model is 6 818 590 and with the discrete logistic model amounts to 10 891 700. This as we see is exceeded by a number of 4 million infected, a number that is really worrying when predicting these values. So we could continue to mention for all the cases of South America, in which they have similar behaviors.

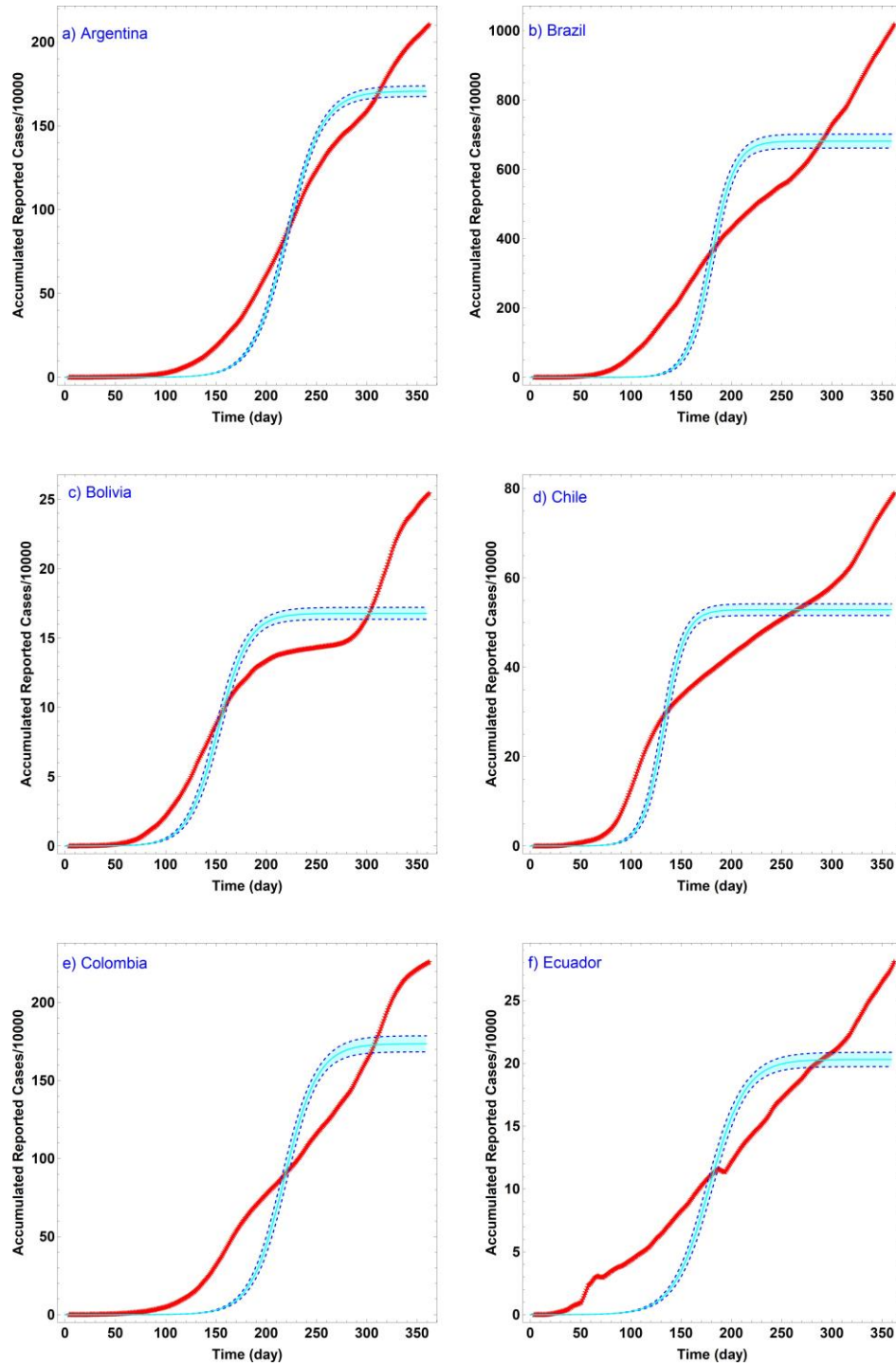


Figure 3.10: Differential logistic model using moving average data with 95% of confidence level. a) Argentina, b) Brazil, c) Bolivia, d) Chile, e) Colombia, f) Ecuador. Total cases data as function of time (red points) and fitting model with bands (blue lines)

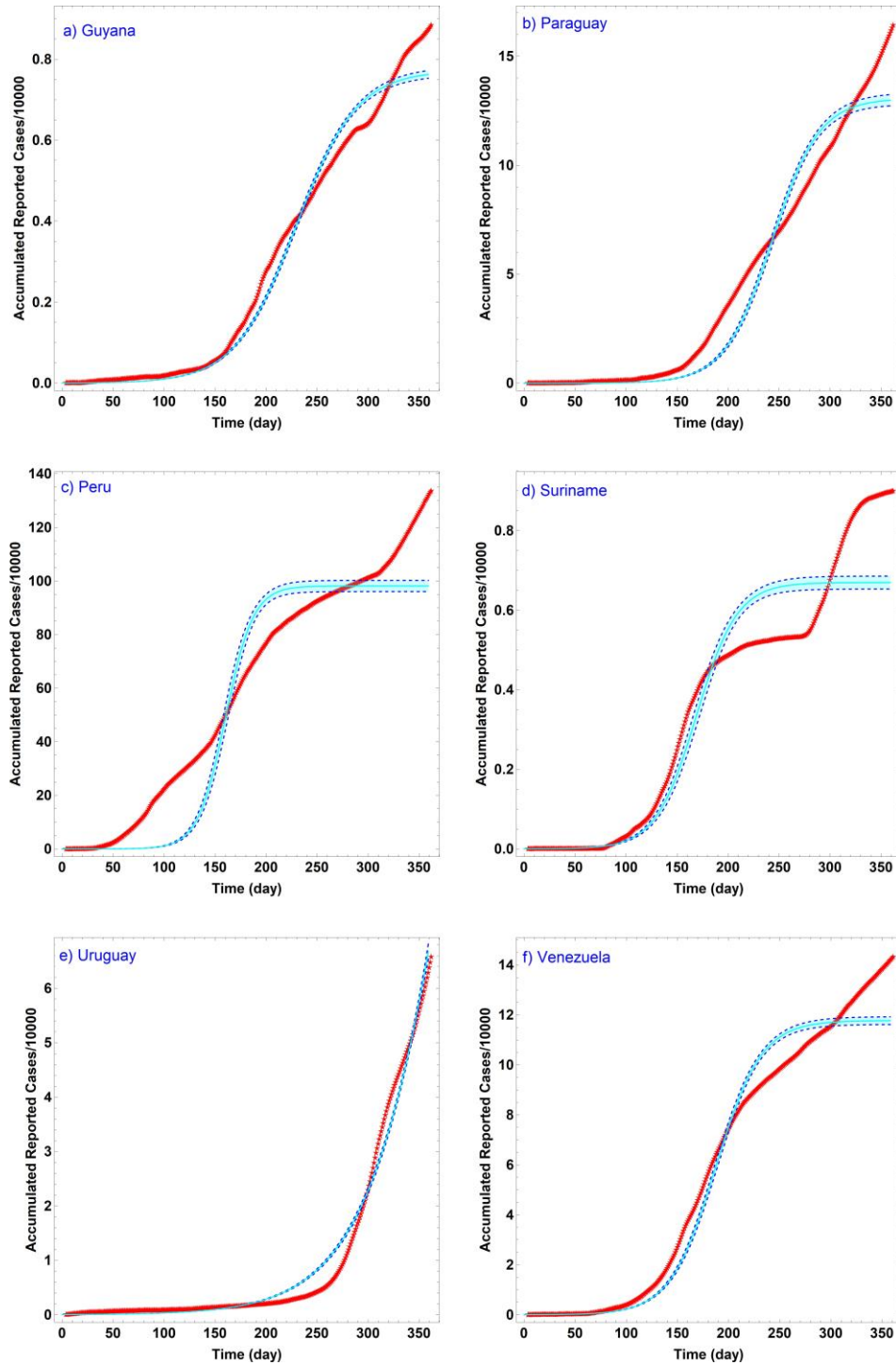


Figure 3.11: Differential logistic model using moving average data with 95% of confidence level. a) Guyana, b)Paraguay, c) Peru, d) Suriname, e) Uruguay, f) Venezuela. Total cases data as function of time(red points) and fitting model with bands (blue lines).

Logistic Model Results					
Country Name	Estimated cases $N_{max}$	Standard error( $\pm N_{max}$ )	Growth rate $\lambda_{log}$	Standard Error( $\pm log$ )	$R^2$
Argentina	1 707 840	15 988.9	0.057	0.001	0.976
Brazil	6 818 590	102 619	0.085	0.001	0.933
Bolivia	167 909	2 174	0.066	0.001	0.951
Chile	528 845	6 599	0.094	0.001	0.951
Colombia	1 736 150	26 201.2	0.060	0.001	0.941
Ecuador	203 199	2 892	0.055	0.001	0.944
Guyana	7 723.1	0.001	0.034	0.001	0.991
Paraguay	130 716	1 359	0.043	0.001	0.978
Peru	9 811 770	10 668.2	0.073	0.001	0.964
Suriname	6 694.56	82.730	0.050	0.001	0.958
Uruguay	265 128	28 342.7	0.021	0.001	0.986
Venezuela	117 764	777.711	0.045	0.001	0.988

Table 3.4: Estimated cases ( $N_{max}$ ) and growth rates ( $\lambda$ ) parameters with their standard errors from the differential logistic model for the different countries under study.

However, the important thing that can be highlighted with the use of this model is undoubtedly the fact that the approximation of the available data are good enough ( $R^2 > 0.8$ ) for the use of all the cases studied in this work. With this in mind, in terms of the growth rate of this model it is necessary to mention that there appears a reduction. Analyzing the growth rate of the differential logistic model with those of the exponential model, we notice that this parameter has decreased at least two or three times less than the found in Equation 2.6. This parameter undergoes a change in the range of 1.52 times until 3.18 times less. This tells us that the model used in this section is useful when we seek to analyze how COVID-19 spread in South America.

### 3.5 Differential Logistic Model in Function of Daily and Accumulated Cases

Based on the last results we evidence that the discrete logistic model and the differential logistic model cannot be used to properly predict the total number of cases when a sigmoidal trend is not observed. With this in mind, a more accurate value of the maximum number of infected people can be predicted by means of Equation 2.11. This equation shows a simple relationship between daily reported cases and total cases but allows calculating the growth rate coefficient and the maximum number of cases with current state policies.

Figures 3.12 and 3.13, show the actual data in red and the estimated data in blue line with their respective confidence intervals (blue light). Moreover, the parameters of the estimated cases and growth rate are shown in Table 3.5. The important thing in these figures is the fact that it shows us a more accurate approximation of when the pandemic state could end. This means that you can more accurately predict the dates and likely values that you will reach your maximum level of infection in the South American population.

The figures obtained with the use of the Equation 2.11, with the accumulated total data and the daily cases, allow us to better observe an approximation of when the values of the curves decrease. In other words, the fitting curve that we achieved shows us on the X axis, the estimated number of total cases accumulated. Although in all the countries analyzed the values are different, they always have a trend very similar to the actual data. This model predicts the peaks of rise of COVID-19 infection in the population. Therefore, it can be used for data analysis both for the year and for dates later. The estimated values obtained, for all cases exceed those that have already been found in the year, that is, it foresees the cases that could continue to occur during the following years. This fact helps us to alert the corresponding authorities to take more rigorous measures in order to prevent a third or fourth wave of infections that could be disastrous throughout South America.

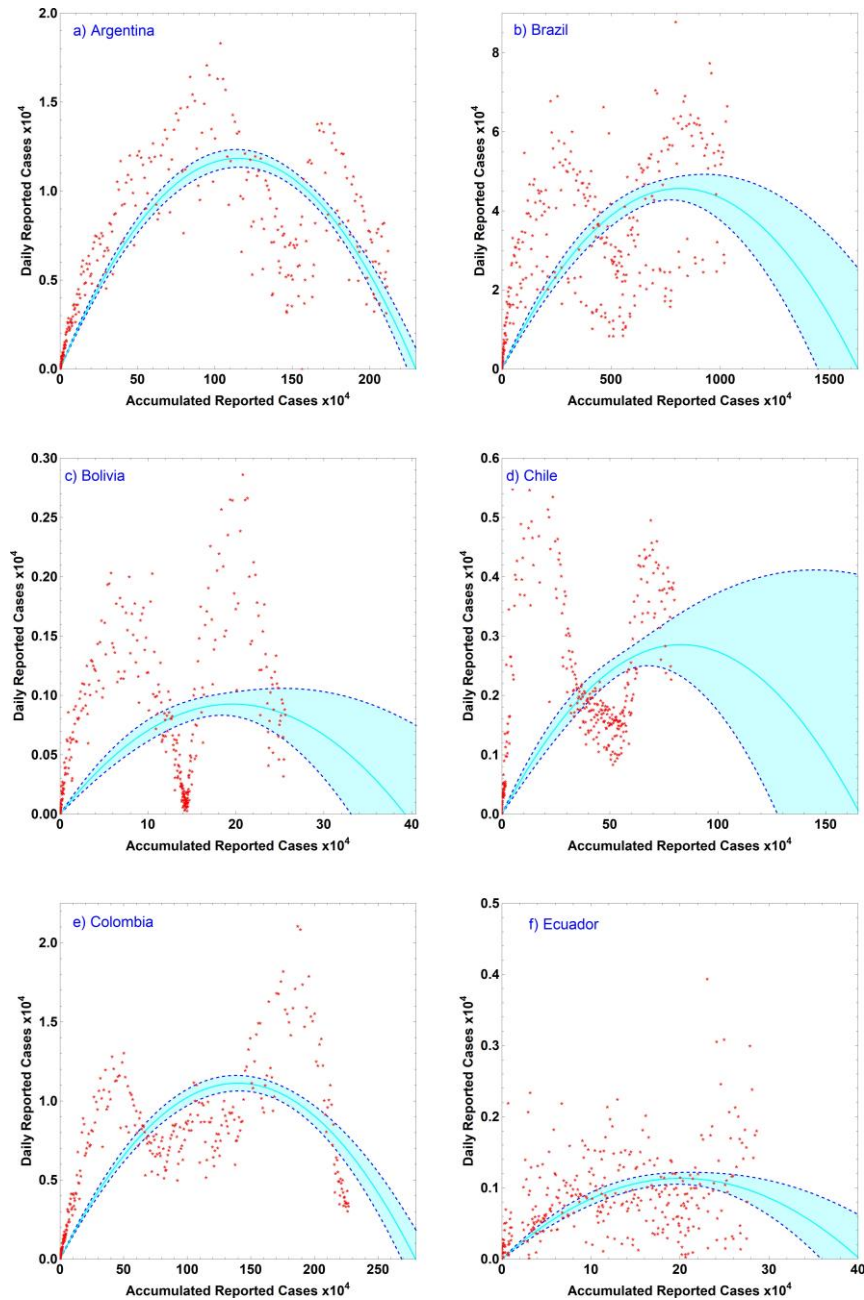


Figure 3.12: Differential logistic model using moving average data with 95% of confidence level. Using the relationship between daily reported cases and total cases. a) Argentina, b) Brazil, c) Bolivia, d) Chile, e) Colombia, f) Ecuador. Total cases vs Daily cases data as function of time (red points) and fitting model with bands (blue lines)

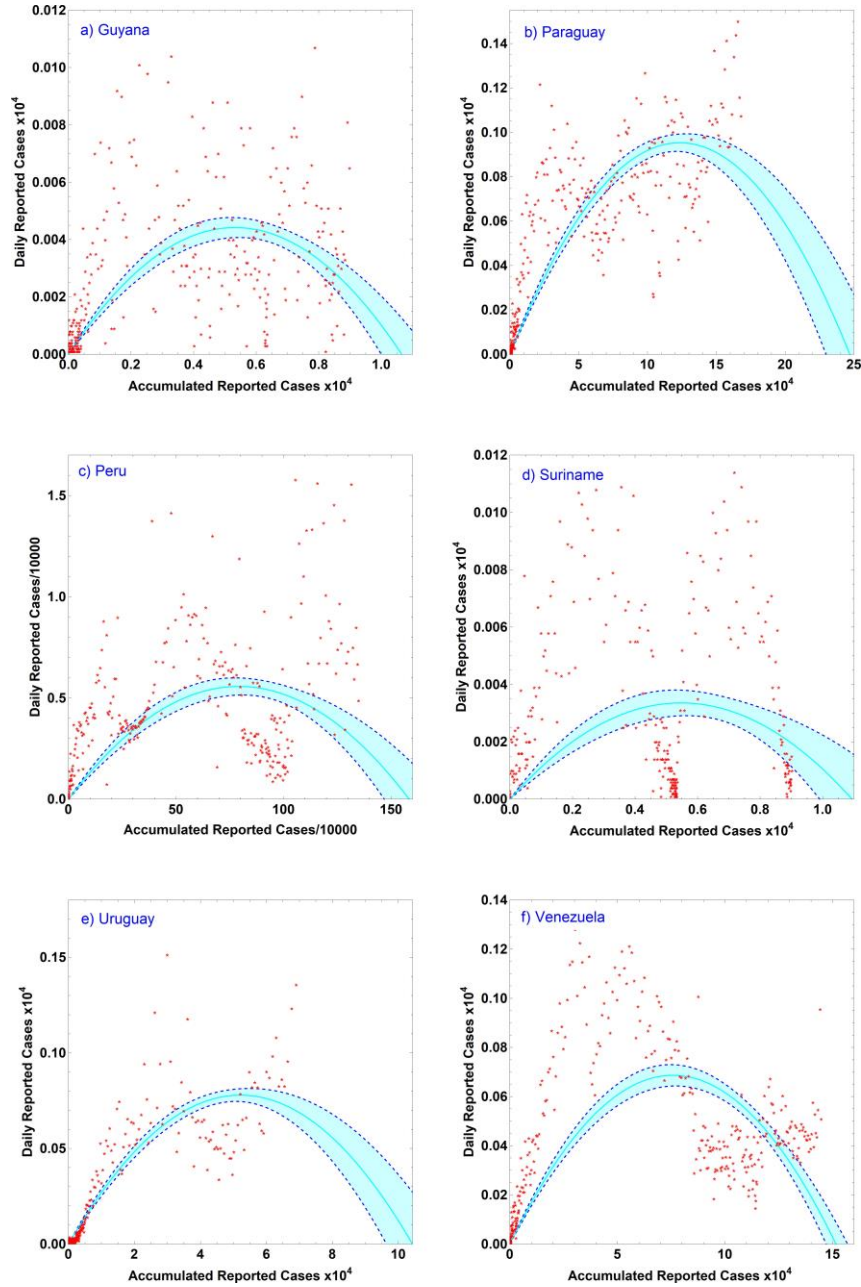


Figure 3.13: Differential logistic model using moving average data with 95% of confidence level. Using the relationship between daily reported cases and total cases. a) Guyana, b) Paraguay, c) Peru, d) Suriname, e) Uruguay, f) Venezuela. Total cases vs Daily cases data as function of time (red points) and fitting model with bands (blue lines).



Logistic Model Results					
Country Name	Estimated cases $N_{max}$	Standard error( $\pm N_{max}$ )	Growth rate $\lambda_{log}$	Standard Error( $\pm log$ )	$R^2$
Argentina	2 297 940	28 987	0.020	0.000	0.87
Brazil	16 308 300	1 158 080	0.011	0.000	0.77
Bolivia	392 396	42 903	0.009	0.000	0.55
Chile	1 659 926	297 800	0.006	0.000	0.71
Colombia	27 947 443	60 094	0.016	0.001	0.84
Ecuador	402 109	27 617	0.011	0.001	0.70
Guyana	10 651	386	0.016	0.001	0.71
Paraguay	246 933	10 082	0.015	0.001	0.87
Peru	1 583 936	69 622	0.014	0.001	0.71
Suriname	10 934	663	0.012	0.001	0.50
Uruguay	104 246	4 625	0.029	0.001	0.87
Venezuela	151 578	2 495	0.018	0.001	0.80

Table 3.5: Estimated cases ( $N_{max}$ ) and growth rates ( $\lambda$ ) parameters with their standard errors from the differential logistic model for the different countries under study based on daily reported cases and total cases.

While the analysis of the growth rate is important in all models, here it is found that it decreases in excessive amounts. Comparing the results of the growth rate with the previous models, exponential model and discrete logistic model, it decreases by about twenty times and five times less, respectively. Additionally, as analyzed, the  $R^2$  in this specific case is  $> 0.7$ . This does not mean that fittings are not inadequate or inefficient, but that greater care must be taken when choosing the values that should be taken both for accumulated total values and for the values of daily cases, except in cases like Suriname and Bolivia, where  $R^2 < 0.7$ . However, in these cases the estimated values are suitable for the analysis of the same as mentioned above.

Percentage Analysis					
Country Name	Total real cases	Estimated cases $N_{max}$ (Discrete Logistic Model)	Relative Error (%)	Estimated cases $N_{max}$ (Differential Logistic Model)	Relative Error (%)
Argentina	2 118 676	2 099 630	0.89	1 707 840	19.39
Brazil	10 390 461	10 891 700	4.82	6 818 590	34.37
Bolivia	256 462	244 260	4.75	167 909	34.52
Chile	803 009	695 250	13.41	528 845	34.14
Colombia	2 269 582	2 701 910	19.04	7 736 150	23.50
Ecuador	286 155	304 050	6.25	203 199	28.98
Guyana	8 993	8 910	0.92	7 723	14.12
Paraguay	168 043	179 680	6.92	130 716	22.21
Peru	1 349 847	1 176 770	12.82	981 177	27.31
Suriname	9 022	9 090	0.75	6 695	25.79
Uruguay	69 074	77 420	12.08	265 128	283.83
Venezuela	144 786	130 870	9.61	117 764	18.66

Table 3.6: Estimated cases ( $N_{max}$ ) and midpoints ( $t_0$ ) parameters with their standard errors from the discrete logistic model for the different countries under study.

### 3.6 COVID-19 Outbreak in Ecuador

For this section we have focused on the specific case of Ecuador. It is the country where we recurred and we can analyze it more depth in terms of social, economic and other aspects. Ecuador is a country located to the northwest of South America, bordering Colombia in the north, Peru to the south and east, and the Pacific Ocean to the west. Ecuador has 24 provinces, among which the main ones are Pichincha and Guayas, With population densities of 333.08 people for  $km^2$  and 275.47 people per  $km^2$  respectively and its capital cities Quito with a population of 1399814 and Guayaquil with a population of 1952029 respectively according to Worldometer. According to Brazilian Reports, in Latin America the city with the highest number of reported cases of COVID-19 was Guayaquil, thus becoming the epicenter in the continent of this infection<sup>39</sup>.

Guayaquil's real situation was much more intense than many media outlets showed it. There are several images showing sick people dying in the streets of the city and the little action taken by the government to stop this situation.

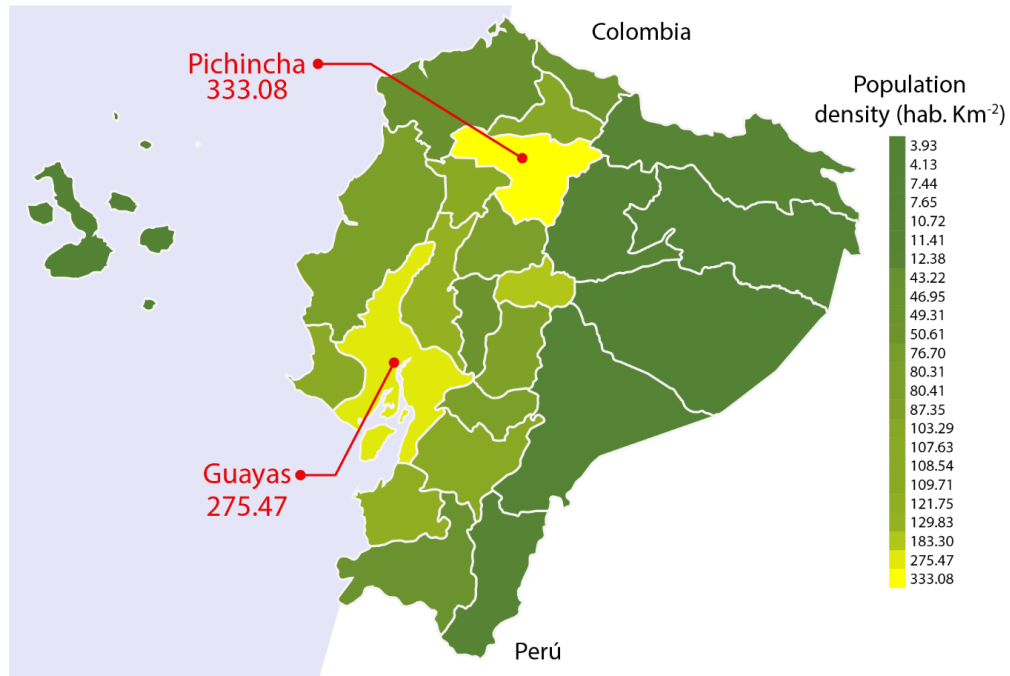


Figure 3.14: Ecuador population density map

The spread of COVID-19 in Ecuador is shown in the Figure 3.15. The data reported are shown in black and the data corrected with the 7 point moving average in red. As mentioned at the beginning of the research, all this is used to avoid peaks or valleys that cause errors at the time of modeling.

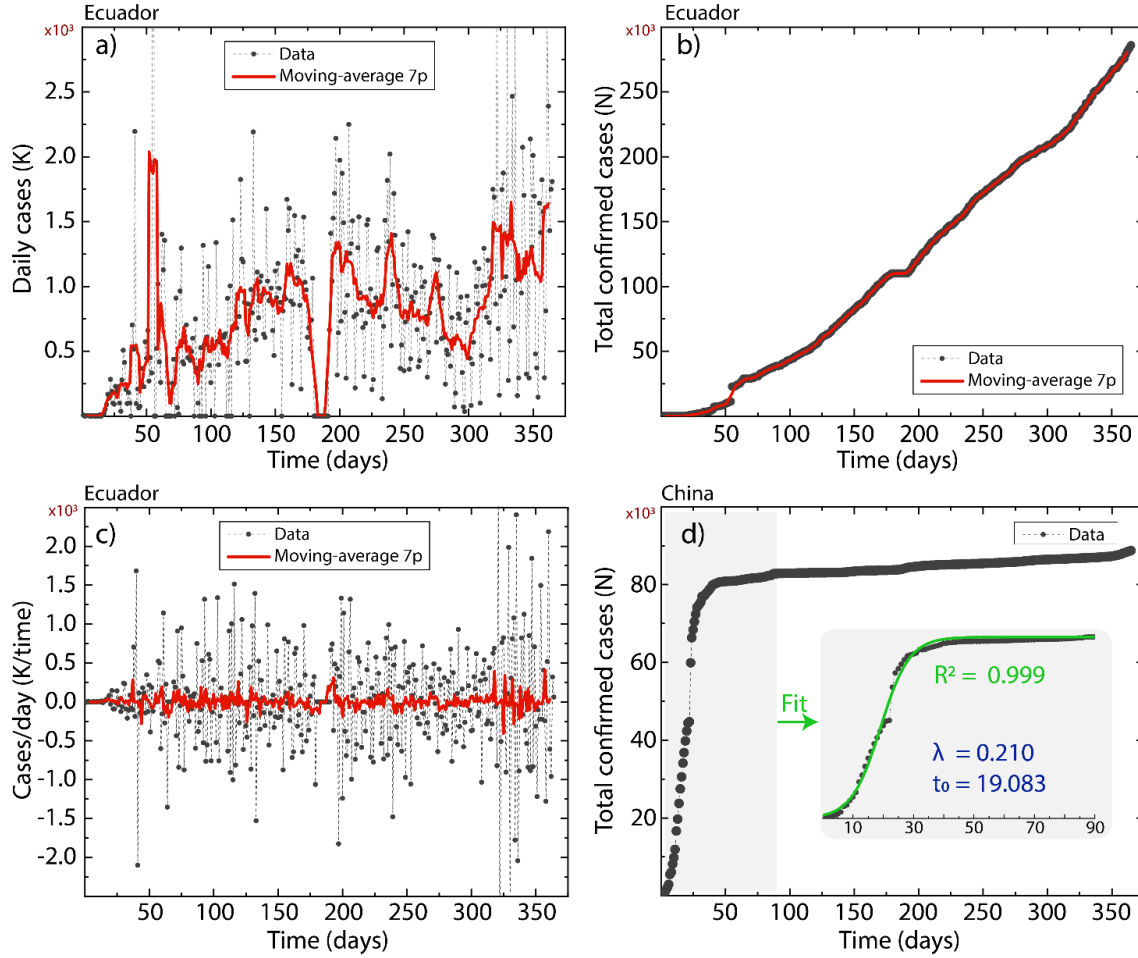


Figure 3.15: Ecuador: a) the number of infected people per day, b) the total number of confirmed cases reported in a 365-days window, and c) the forward differential method applied to the daily reported cases. China: d) the total number of confirmed cases reported on 12/31/2020. Data are represented in black and the 7-point average data in red. The inset in Figure 1d shows the fit of China data using the discrete logistic model.

Figure 3.15a shows the daily cases reported as a function of time with a 365-day window. As can be seen the reported data are characterized by intense peaks exceeding  $1.5 \times 10^3$  cases per day. On the other hand, within the data we can observe a gap related to days in which there were zero reported cases. This can be seen graphically as a deep valley between the days 180 and 190

approximately from August 27 to September 7, 2020. As you can see the scattering of the 7-point moving average data is remarkable in comparison to the reported data. However, these maintain the same trend and can be used for the complete analysis of the models.

Figure 3.15b shows the total number of cumulative cases reported as a function of time. Here there is a small interruption because non new cases were reported on the dates of May 4 to May 8, 2020 approximately. This is once again corrected by the 7-point moving average we made earlier. Figure 3.15d has also been added which shows the same data but in the case of China, this to make a comparison between these two countries. In the case of China, an early saturation is shown in the first 100 days from the first confirmed case, which contrasts with the case of Ecuador where recreation grows exponentially day after day despite having had more than two months to prepare against the COVID-19.

In terms of identification of waves of contagion, Figure 3.15c has been made showing the use of the forward differential method applied to the daily case data reported. However, it is not possible to observe or differentiate between the first, second or third wave, because the data remains in a constant trend of  $\pm 1.0 \times 10^3$ . This suggests that the population is in constant contagion due to the few controls and preventions that exist in the country.

If we put together the models made previously with respect to Ecuador we can obtain a graphical comparison and we can also make a table in which the results of each proposed model are better visualized. Figure 3.16 shows the curves of the different models proposed: the exponential model (Figure 3.16a), the discrete logistic model (Figure 3.16b), and the solution of the differential logistic model (Figure 3.16c). Moreover, the estimated values for the constants and for the evaluated parameters are reported in Table 3.8.

As a first observation in the results table, we analyzed the growth rate coefficient estimated of the exponential model ( $\lambda = 0.185 \text{ day}^{-1}$ ), which is about 13.21 times larger than that found by the discrete logistic model ( $\lambda = 0.014 \text{ day}^{-1}$ ) and 3.36 times larger than that found by the differential logistic model ( $\lambda = 0.055 \text{ day}^{-1}$ ). The best explanation for this high growth rate is due to the little or no preventive action at the beginning of the pandemic.

The BBC shows that Ecuador was in the second position of the countries most affected and with the highest number of infected by COVID-19 in all Latin America only after Brazil<sup>40</sup>. Additionally, The New York Times reported on April 23, 2020 that the total death toll was 15 times more than the government intended to show<sup>41</sup>. Guayaquil is one of the cities with the

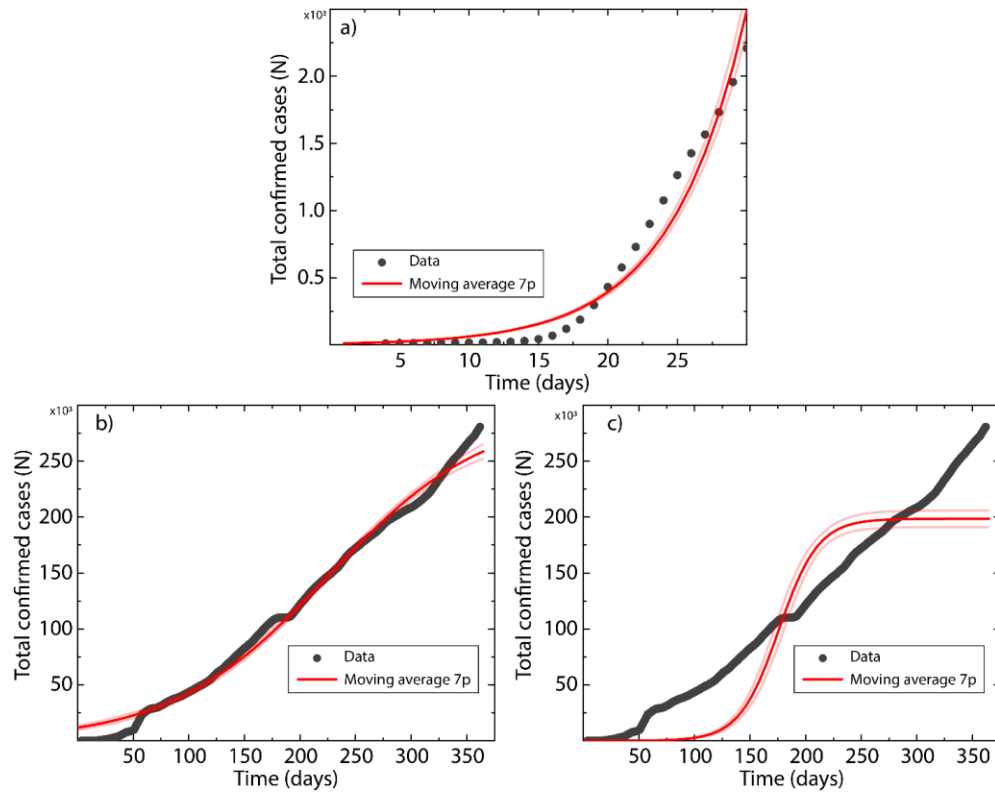


Figure 3.16: The total number of cases as a function of time fitted with different models: a) exponential model (Equation 2.6), b) discrete logistic equation (Equation 2.7), and c) solved ODE logistic equation (Equation 2.9). Data and fit are represented in black and red, respectively. Pink lines represent the confidence level (%95)

<b>Exponential Model-Equation 2.6</b>			
<b>Parameter</b>	<b>Estimated</b>	<b>Standard error</b>	<b>R<sup>2</sup></b>
$\lambda$	0.185	0.001	0.978
<b>Discrete logistic Model-Equation 2.7</b>			
$\lambda$	0.014	0.000	0.997
$t_0$	230.566	2.4333	
$N_{max}$	304 050	3 960	
<b>Differential Logistic Model-Equation 2.9</b>			
$\lambda$	0.055	0.001	0.944
$N_{max}$	203 199	2 892	

Table 3.7: Estimated parameters using the different models of the present work. The  $\lambda$  represents the estimated (exponential or logistic) growth rate,  $N_{max}$  is the maximum number of infected people, and  $t_0$  denotes the midpoint.

highest number of deaths both in hospitals, in their own homes and in more extreme cases in the streets of the city. This was due to the fact that only a partial quarantine was implemented as of March 17, 2020, when the first suspicion of the first case was already reported as of February 14, 2020.

Analyzing Figure 3.16a, we found a clear exponential growth within the first 30 days of the pandemic with some points below or above the exponential fit (red line) and their confidence intervals (pink lines). In addition, within 30-days window your  $N_{max} = 2240$  until March 31, 2020. On the other hand, the discrete logistic model shows an excellent coincidence between the reported data and the fitting curve (Figure 3.16b). In terms of comparison between Ecuador and China, we can find that its midpoints of both Ecuador ( $t_0 \sim 230$ ) and China ( $t_0 \sim 19$ , Figure 3.16d), are extremely different about 12 times greater than that of the Asian country, this tells us that the pandemic was managed within the first month. In contrast, we can mention that in Ecuador the flattening of the infection curve began on October 12, 2020, that is, more than 7 months after the pandemic arrived. Another point at which we should emphasize is in Figure 3.16c that shows us a discrepancy between the reported data and the fitting curve, leaving us with the suggestion that modeling should be used by the logistic differential method with extreme caution when trying to predict long-term values. Specifically, the total number of infected persons

calculated by Equation 2.9 ( $N_{max} = 203199$ ), which has a difference of at least 1000000 total cases compared to what was estimated by the discrete logistic model (Equation 2.7) ( $N_{max} = 304050$ ). Despite this difference, the values obtained in both models have a parameter of  $R^2$  above that it is taken as sufficient ( $> 0.8$ ). This data is confirmed by Table 3.8, in which both models have  $R^2 > 0.9$ , which allows us to suggest that both models fulfill their purpose and fits for the data of Ecuador taken during the year.

To analyze the growth rate coefficient of each model in a qualitative way, it has been proposed to represent them in the form of a matrix (Figure 3.17). In this way we can analyze how the growth was and in turn how health policies (social distancing, testing, quarantine, treatment of infected patients) were applied to reduce it in Ecuador. Figure 3.17 shows that the growth rate coefficient suffers a decrease that we have already mentioned also the days where the contacts of infected people were given. The analysis of the growth rate coefficient of the exponential model (Figure 3.17a), the discrete logistic model (Figure 3.17b) and the differential logistic model (Figure 3.17c), we realize that they are reduced to around  $\sim 92\%$  and  $\sim 70\%$  respectively. This fact confirms that there was a behavior of the chaotic or random type as a function of time. Once the pandemic began, the policies implemented greatly affect the growth rate coefficient since these were taken recklessly, without considering all the effects they could have on citizens and their health system. To this is added, that the time in which a complete quarantine was implemented throughout the country was not enough and the population was released with the only rule of maintaining a social distancing. Finally, it is that in that year once the state of emergency and quarantine ended, the celebration of several festivals and events that had no controls by the government or the ministry of public health was allowed.



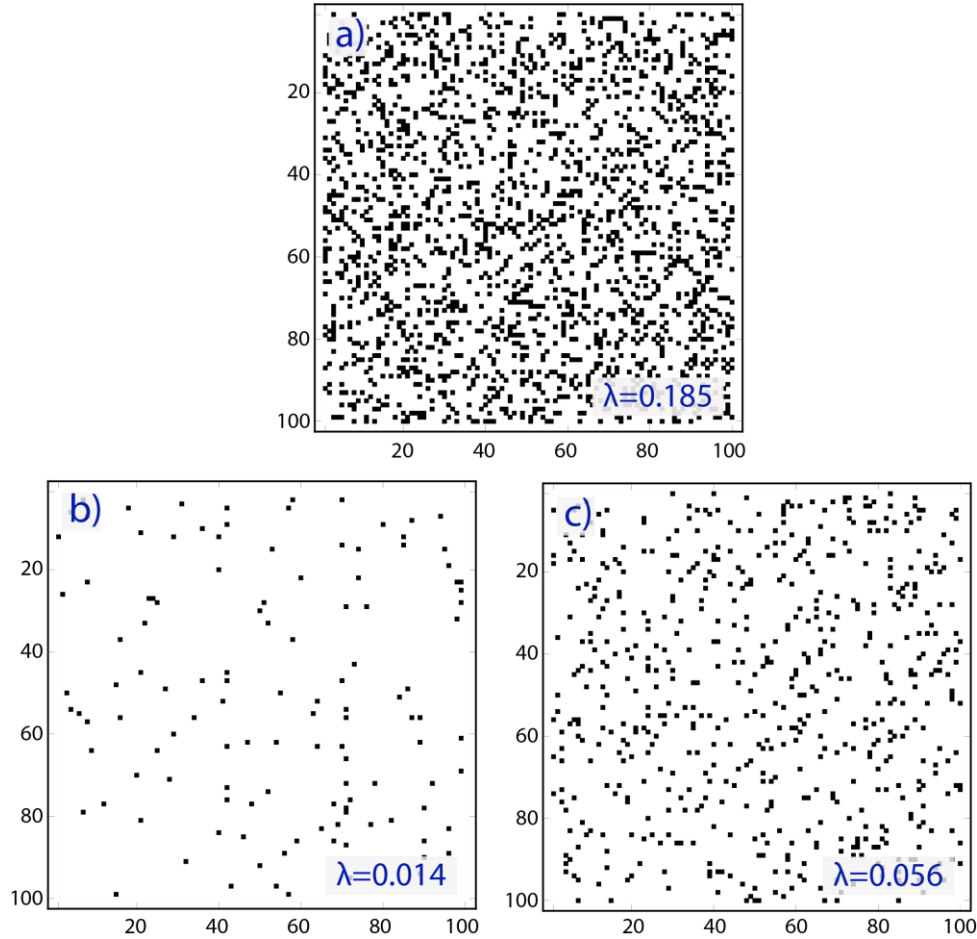


Figure 3.17: Matrix representation of the growth rate coefficient ( $\lambda$ ) from (a) exponential model (for simplicity is taken), (b) discrete logistic equation, and (c) differential logistic model. Black markers represent the infected people affected by COVID-19.

It is important to analyze more deeply the total case approximation ( $N_{max}$ ) of each model used. As we have seen, the discrete logistic model (Equation 2.7) gives us an  $N_{max} = 304050$  and the differential logistic model (Equation 2.9) an  $N_{max} = 203199$ . However, the actual values already exceed these approximations on dates of March 17, 2021 with  $N_{max} = 305598$  and December 17, 2020 with  $N_{max} = 204249$ , respectively. This tells us that the models cannot be used correctly to predict a total number of infected cases in Ecuador, because, they do not have sigmoid behavior,

as can be seen in Figure 3.15d. Given this case, we have set out to analyze it in a different way using Equation 2.11. Although, in the previous section we have analyzed in a general way the use of this equation for different countries. Here, we want to emphasize in a comparison with the countries that most closely resemble the behavior of their curves. The selected countries are Brazil, Peru and Colombia. We have put the data for these countries in Table 3.7, for better visualization at the time of comparison. At the same time we change the way we represent the model to differentiate them from the models we already present. With Equation 2.11 we obtained parabolic approximation curves (red color) that are in accordance with the data reported for each country ( $R^2 > 0.8$ ), having as much as  $N_{max} = 402109$  for Ecuador,  $N_{max} = 3.923 \times 10^5$  for Brazil,  $N_{max} = 1,583 \times 10^6$  for Peru, and  $N_{max} = 2,794 \times 10^6$  for Colombia.

The estimated  $N_{max}$  value in Ecuador agrees with the data reported at the end of May 2021 ( $N_{max} = 426037$ ) (see the following link for the real-time updated map on the Ecuador situation of cases of coronavirus: <https://www.worldometers.info/coronavirus/country/ecuador/>). This suggests that the proposed logistic model may be useful and further calibrated if more data from daily cases are added.

Analyzing the parameters obtained for Ecuador, we found that the growth rate estimated ( $\lambda = 0,011 \text{ day}^{-1}$ ) is reduced by at least  $\sim 94\%$  compared to the exponential model ( $\lambda = 0,185 \text{ day}^{-1}$ ). Moreover, this resonates obtained with Equation 2.11, has a great similarity with that reported by the discretionary logistic model ( $\lambda = 0,014 \text{ day}^{-1}$ ). Curiously, they are also intimately linked to those found in the other countries Brazil ( $\lambda = 0,001 \text{ day}^{-1}$ ), Peru ( $\lambda = 0.0013 \text{ day}^{-1}$ ) and Colombia ( $\lambda = 0,016 \text{ day}^{-1}$ ), confirming to us that they had a similar management by the authorities and the population, at the time of fighting against the rapid spread of COVID-19. It is necessary to mention that scatter data points are observed with respect to the parabola. Telling us that the variability of the growth rate coefficient must be considered in a mandatory way as a function of time, mainly due to the fact that there are still new cases reported on a daily basis. We have defined the coefficient variability, which can be calculated with the use of Equation 2.12 and which can be expressed either as a function of the total cases and also as a function of the time.

As we have focused on studying the specific case of Ecuador, Figure 3.19 shows us how the  $\lambda$  coefficient varies. The information obtained tells us that in the first  $10 \times 10^3$  confirmed cases, there is an extremely high growth rate ( $0.2 < \lambda < 0.08$ ). After this, in the  $20 \times 10^3$  and  $30 \times 10^3$  total confirmed cases, this value decreases to almost half (0.05) and to an eighth part  $\sim 0.01$ ,

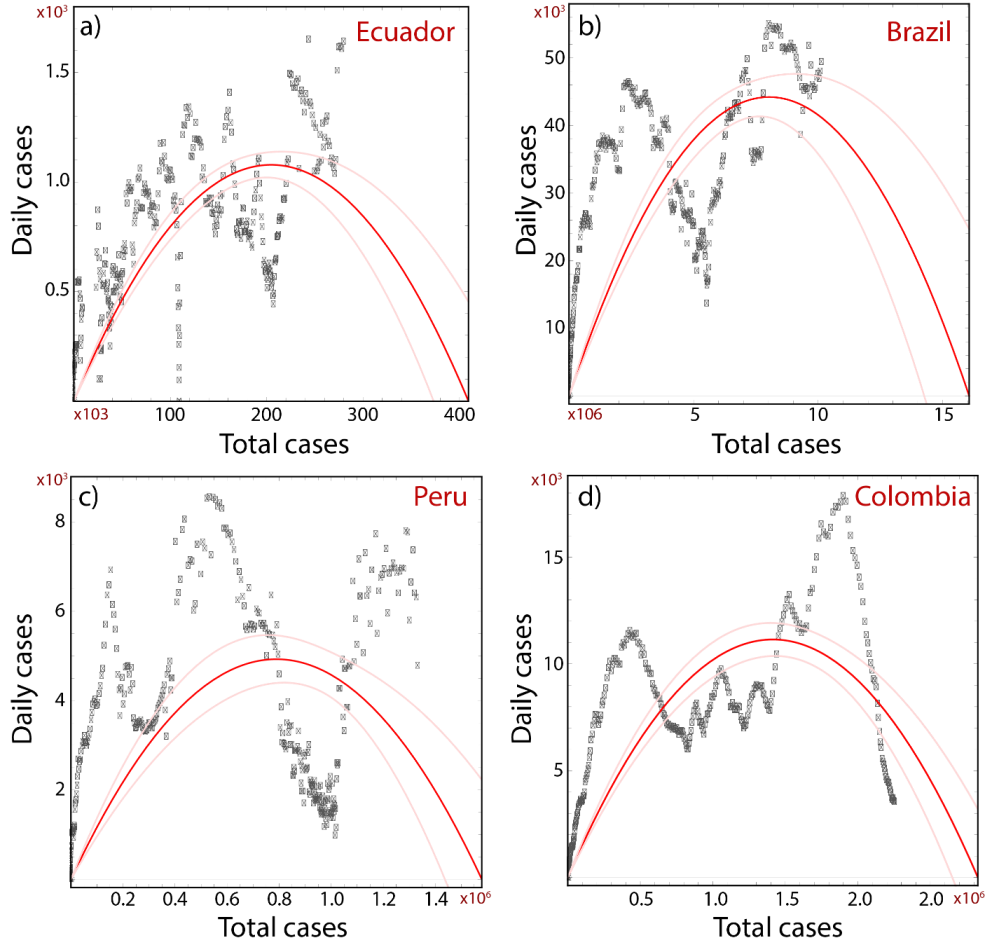


Figure 3.18: The daily cases as a function of the total confirmed cases. The black markers show the data and the red curve is the regression according to Equation 2.11. Pink lines represent the confidence level (95%). For comparison, it is analyzed different countries: a) Ecuador, b) Brazil, c) Peru, and d) Colombia.

respectively. Our growth rate, takes its value when the number of total cases confirmed amounts from  $40 \times 10^3$  to  $300 \times 10^3$ , in a range of  $0.02 < \lambda < 0.005$ . Interestingly, we can define about three important peaks where you can see the increase in cases. Based on Figure 3.19c, the first peak is between  $80 \times 10^3 - 160 \times 10^3$ , the second peak between  $170 \times 10^3 - 280 \times 10^3$ , and the third

<b>Equation 2.11</b>			
<b>Parameter</b>	<b>Estimated</b>	<b>Standard error</b>	<b>R<sup>2</sup></b>
<b>Ecuador</b>			
$\lambda$	0.011	0.001	0.902
$N_{max}$	402 109	2 761	
<b>Brazil</b>			
$\lambda$	0.011	0.001	0.55
$N_{max}$	$3.923 \times 10^5$	$4.290 \times 10^4$	
<b>Peru</b>			
$\lambda$	0.013	0.001	0.71
$N_{max}$	$1.583 \times 10^6$	$6.96 \times 10^4$	
<b>Colombia</b>			
$\lambda$	0.016	0.001	0.84
$N_{max}$	$2.794 \times 10^6$	$6.0 \times 10^4$	

Table 3.8: Estimated parameters using the different models of the present work. The  $\lambda$  represents the estimated (exponential or logistic) growth rate,  $N_{max}$  is the maximum number of infected people, and  $t_0$  denotes the midpoint.

peak from  $290 \times 10^3$ . Hence the fact of relating these peaks, as a multi-stage behavior COVID-19 outbreak in Ecuador, but not as just simple constant waves of infection. This multi-stage behavior is confirmed by data and celebrations that occurred throughout the year. In terms of dates, our first peak is in the week of September 28 to October 4, 2020, and the second peak coincidentally is from December 7 to 13, 2020, dates where Christmas and New Year is celebrated.

Adding, Figure 3.19b, has an importance in analyzing the symmetry of the growth rate coefficient distribution. This figure has a slight skewed on its right, confirming that this distribution is asymmetrical, where there are points that the growth rates are larger than others. It is important to mention that the largest number of growth rates are located below  $0.02 \text{ day}^{-1}$ . If we focus on this section (Figure 3.19d), we can observe an symmetrical behavior. Additionally, make it known that most growth rate coefficients are between  $0.007$  and  $0.017 \text{ days}^{-1}$  and the frequency with the highest value at  $0.001 \text{ day}^{-1}$ .

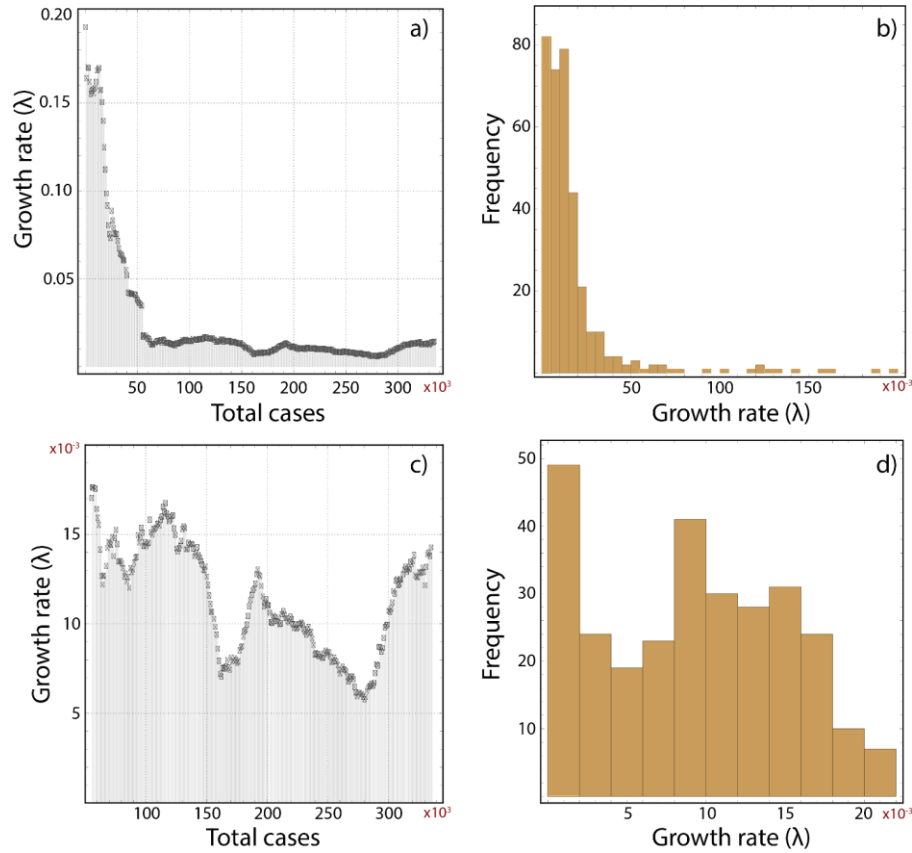


Figure 3.19: a), c) The growth rate coefficient as a function of the total confirmed cases computed using Equation 2.12 and considering the data in a one-year window for Ecuador. b), d) the associated histograms.

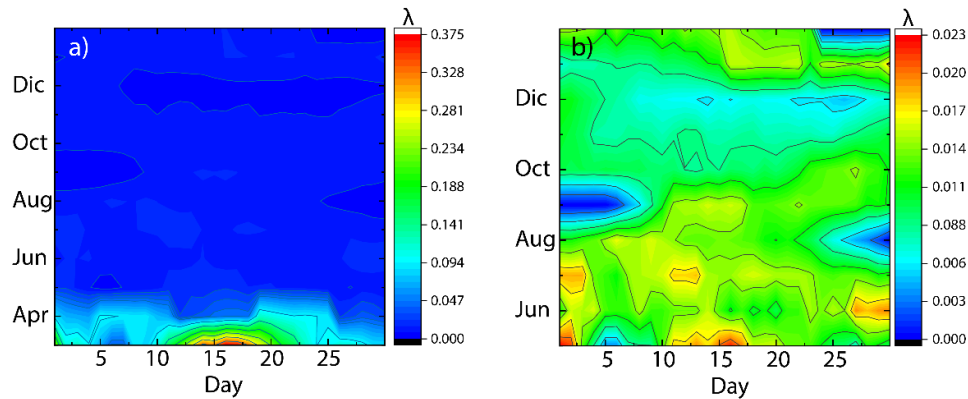


Figure 3.20: The growth rate coefficient of Equation 2.12 is represented as a surface-color plot vs 30-days and 12-months. (a) the growth rate in a one-year window and (b) the growth rate from May 1, 2020, to February 28, 2021.

Finally, Figure 3.20 displays the growth rate coefficient as a function of time. Mainly, the critical months of Ecuador were located in March and April 2020, specifically in the third week, where the values of the coefficient are the maximum found (Figure 2.20a). The variability of the coefficient in the following months (May 2020 to February 2021) is shown in Figure 2.20b, where values are not extremely high as the first months. In fact, this fact is attributed to the care and protection that society adopted with the use of face masks, alcohol, disinfectant gel and the decision itself to mobilize as little as possible.



## Chapter 4

### Conclusions & Outlook

We have presented a simple mathematical approach based on the conventional exponential model, the discrete logistic equation, and the differential logistic model to scrutinize the COVID-19 outbreak in South America. The estimated parameters were the growth rate ( $\lambda$ ), the total number of cases ( $N_{max}$ ), and the midpoint of the maximum growth ( $t_0$ ). The one year data from the first day of reported cases of each country, mostly from February, 2020 to March, 2021. The real data was subdue to a 7-point moving average to avoid some problems(zero days reported cases) during the modeling. As a first approach, it is demonstrated that the estimated growth rate in every case suffers a decrease in the value calculated. Following a rule from highest to lowest value, we can say that growth rate calculated by Exponential model > Discrete Logistic Model > Differential Logistic Model. Despite the fact that there was a lockdown in each country, affecting not only the economic status, the collapse of the healthy systems had to happen, due to the number of infected people increases exponentially at the first 30 days. In general way, when we compare Discrete, Differential and the Corrected Equations of Logistic Models, the way to estimated the  $N_{max}$ , not always is correct. Although, the models had a highest approximation, in some cases the  $N_{max}$  were reached before the year were complete. However, this problem can be solved by analyzing daily cases as a function of total cases. Thus, the predicted numerical value of the total number of infected people was found to be in good agreement with the real-time data reported at the end of May 2021. A nicely point at the moment to evaluate each model for each country, is the fact that the fitting curves, were acceptable ( $R^2 > 0.8$ ).



Here, we also emphasize the case of the COVID-19 outbreak in Ecuador. Selecting some countries with almost same behavior like Brazil, Peru and Colombia. Additionally, for this specific case we analyze the variability of the growth rate coefficient a function of the total cases and time. In context, the results evidenced as we mention that the growth rate decreased from  $=0.185 \text{ day}^{-1}$  (exponential model) to  $=0.014 \text{ day}^{-1}$  (discrete model) and  $=0.056 \text{ day}^{-1}$  (differential model) over the months. During the firstly days of pandemic, as in the other countries a lockdown happens in Ecuador, guide to reduce the fast spread of the infection. This policy had a “relative” positive effect in terms to reduce the growth rate values. We can corroborated this fact by analyzing the growth rate variability as function of time. Mainly it is observed that the most critical months of the pandemic were March and April 2020. As we conclude before, discrete ( $N_{max} \sim 305k$ ) and differential( $N_{max} \sim 204k$ ) estimated total cases, are not allowed to be used in a approximation cases, because both can be underestimated. This last simple mathematical approach (Variability Growth rate) for the Ecuador case also can be applied in other South American countries mentioned here to take immediate action against the COVID-19 disease.

# Appendix A

## Long Appendix 1 Heading

### A.1 Exponential function calculus

Equation may be written

$$\frac{dN}{dt} = kN \quad (\text{A.1})$$

The solution is of the form

$$N = N_0 a^{bt} \quad (\text{A.2})$$

where  $a$  and  $b$  are constants. The time derivate of  $N$  is

$$\frac{dN}{dt} = b (\ln a) N_0 a^{bt} = b (\ln a) N \quad (\text{A.3})$$

This satisfies Eq.(A.1) provided

$$k = b \ln a \quad (\text{A.4})$$

we must choose values to let  $a = e$  where

$$e = \lim_{x \rightarrow \infty} 1 + \frac{1}{x} = 2.718.... \quad (\text{A.5})$$

then  $k = b \ln e = b$  and the solution to Eq.(A.1) is

$$N = N_0 e^{kt} \quad (\text{A.6})$$

## A.2 Solution for ODE differential logistic equation

Rewrite the differential form of the logistic equation (Equation 2.8)

$$\frac{dN(t)}{dt} = \frac{\lambda N(t)(N_{max} - N(t))}{N_{max}} \quad (A.7)$$

Then multiply both sides by  $dt$  and divide both sides by  $N(t)(N_{max} - N(t))$

$$\frac{dN(t)}{N(t)(N_{max} - N(t))} = \frac{\lambda}{N_{max}} dt \quad (A.8)$$

Multiply both sides of the equation by  $N_{max}$  and integrate:

$$\int \frac{N_{max}}{N(t)(N_{max} - N(t))} dN(t) = \int \lambda dt \quad (A.9)$$

The left-hand side can be integrated using partial fraction decomposition. Then the Equation A.9 becomes

$$\int \frac{1}{N(t)} + \frac{1}{(N_{max} - N(t))} dN(t) = \int \lambda dt \quad (A.10)$$

$$\ln |N(t)| - \ln |N_{max} - N(t)| = \lambda t + C \quad (A.11)$$

$$\ln \frac{N(t)}{N_{max} - N(t)} = \lambda t + C \quad (A.12)$$

Now exponentiate both sides of the equation to eliminate the natural logarithm:

$$e^{\ln \frac{N(t)}{N_{max} - N(t)}} = e^{\lambda t + C} \quad (A.13)$$

$$\frac{N(t)}{N_{max} - N(t)} = e^{\lambda t} e^C \quad (A.14)$$

We define  $C_1 = e^C$  so the equation becomes:

$$\frac{N(t)}{N_{max} - N(t)} = C_1 e^{\lambda t} \quad (A.15)$$

To solve this equation for  $N(t)$ :

$$\begin{aligned}
N(t) &= C_1 e^{\lambda t} (N_{max} - N(t)) \\
&= C_1 N_{max} e^{\lambda t} - C_1 N(t) e^{\lambda t} \\
N(t) + C_1 N(t) e^{\lambda t} &= C_1 N_{max} e^{\lambda t} \\
N(t)(1 + C_1 e^{\lambda t}) &= C_1 N_{max} e^{\lambda t} \\
N(t) &= \frac{C_1 N_{max} e^{\lambda t}}{1 + C_1 e^{\lambda t}}
\end{aligned}$$

The last step is to determine the value of  $C_1$ . Substitute  $t = 0$  and  $N_0$  in place of  $N(t)$  in Equation and solve for  $C_1$ :

$$\begin{aligned}
\frac{N(t)}{N_{max} - N(t)} &= C_1 e^{\lambda t} \\
\frac{N_0}{N_{max} - N_0} &= C_1 e^{\lambda(0)} \\
C_1 &= \frac{N_0}{N_{max} - N_0}
\end{aligned}$$

Finally, substitute the expression for  $C_1$  into Equation A.15:

$$N(t) = \frac{C_1 N_{max} e^{\lambda t}}{1 + C_1 e^{\lambda t}} = \frac{\frac{N_0}{N_{max} - N_0} N_{max} e^{\lambda t}}{1 + \frac{N_0}{N_{max} - N_0} e^{\lambda t}} \quad (\text{A.16})$$

Now multiply the numerator and denominator of the right-hand side by  $(N_{max} - N_0)$  and simplify:

$$\begin{aligned}
N(t) &= \frac{\frac{N_0}{N_{max} - N_0} N_{max} e^{\lambda t}}{1 + \frac{N_0}{N_{max} - N_0} e^{\lambda t}} \\
&= \frac{\frac{N_0}{N_{max} - N_0} N_{max} e^{\lambda t}}{\frac{N_{max} - N_0}{N_{max} - N_0} + \frac{N_0}{N_{max} - N_0} e^{\lambda t}} \cdot \frac{N_{max} - N_0}{N_{max} - N_0} \\
&= \frac{N_0 N_{max} e^{\lambda t}}{(N_{max} - N_0) + N_0 e^{\lambda t}} \quad (\text{A.17})
\end{aligned}$$

$$N(t) = \frac{N_0 N_{max} e^{\lambda_{log} t}}{N_{max} + N_0 e^{\lambda_{log} t} - 1} \quad (\text{A.18})$$



# Bibliography

- [1] Shen, C. Y. Logistic growth modelling of COVID-19 proliferation in China and its international implications. *International Journal of Infectious Diseases* **2020**, 96, 582–589.
- [2] Aviv-Sharon, E.; Aharoni, A. Generalized logistic growth modeling of the COVID-19 pandemic in Asia. *Infectious Disease Modelling* **2020**, 5, 502–509.
- [3] Wang, P.; Zheng, X.; Li, J.; Zhu, B. Prediction of epidemic trends in COVID-19 with logistic model and machine learning technics. *Chaos, Solitons & Fractals* **2020**, 139, 110058.
- [4] Ahmed, A.; Salam, B.; Mohammad, M.; Akgul, A.; Khoshnaw, S. H. Analysis coronavirus disease (COVID-19) model using numerical approaches and logistic model. *Aims Bioengineering* **2020**, 7, 130–146.
- [5] Hannah Ritchie, D. B. E. M. J. H. B. M. C. G. C. A. L. R.-G., Esteban Ortiz-Ospina; Roser, M. Coronavirus Pandemic (COVID-19). *Our World in Data* **2020**, <https://ourworldindata.org/coronavirus>.
- [6] South America. <https://www.britannica.com/place/South-America>.
- [7] Satyamurty, P.; Nobre, C. A.; Dias, P. L. S. *Meteorology of the southern hemisphere*; Springer, 1998; pp 119–139.
- [8] OECD;; Bank, T. W. *Panorama de la Salud: Latinoamérica y el Caribe 2020*; 2020; p 154.
- [9] Salud con lupa, F. T.; Martínez, J. Latinoamérica en Cuidados Intensivos. 2020; <https://saludconlupa.com/series/coronavirus/latinoamerica-en-cuidados-intensivos/>.

- 
- [10] Wallace, A. Coronavirus: qué capacidad tienen realmente los países de América Latina para hacer frente a la epidemia de covid-19. <https://www.bbc.com/mundo/noticias-america-latina-51916767>.
- [11] Fowler, A. C.; Fowler, A. C.; Fowler, A. *Mathematical models in the applied sciences*; Cambridge University Press, 1997; Vol. 17.
- [12] Rodríguez Velázquez, J. A.; Steegmann Pascual, C. Modelos matemáticos. **2013**,
- [13] Sodré, U. Modelos matemáticos. *Londrina: UEL* **2007**,
- [14] Vargas, H. E. S.; Sánchez, L. B. R.; Llanes, P. Á. G.; Rodríguez, A. S. Modelación físico-matemática para la toma de decisiones frente a la COVID-19 en Cuba. *Retos de la Dirección* **2020**, *14*, 55–86.
- [15] Meerschaert, M. *Mathematical modeling*; Academic press, 2013.
- [16] Montesinos-López, O. A.; Hernández-Suárez, C. M. Modelos matemáticos para enfermedades infecciosas. *Salud pública de México* **2007**, *49*, 218–226.
- [17] Hernández, J. X. V. Modelos matemáticos en epidemiología: enfoques y alcances. **2007**,
- [18] Martcheva, M. *An introduction to mathematical epidemiology*; Springer, 2015; Vol. 61.
- [19] Brauer, F.; Castillo-Chavez, C.; Feng, Z. *Mathematical models in epidemiology*; Springer, 2019; Vol. 32.
- [20] Adiga, A.; Dubhashi, D.; Lewis, B.; Marathe, M.; Venkatramanan, S.; Vullikanti, A. Mathematical models for covid-19 pandemic: a comparative analysis. *Journal of the Indian Institute of Science* **2020**, 1–15.
- [21] Wang, J. Mathematical models for COVID-19: applications, limitations, and potentials. *Journal of public health and emergency* **2020**, *4*.
- [22] Lial, M. L.; Miller, C. D. *Essential calculus: with applications in business, biology, and behavioral sciences*; Pearson Scott Foresman, 1975.

- [23] Tague, J.; Beheshti, J.; Rees-Potter, L. The law of exponential growth: evidence, implications and forecasts. **1981**,
- [24] Bartlett, A. A. The exponential function—Part 1. *The Physics Teacher* **1976**, *14*, 393–401.
- [25] Allendoerfer, C. B.; Oakley, C. O.; Kerr, D. R. *Elementary functions*; McGraw-Hill Companies, 1977.
- [26] Mousel, S. A. The Exponential Function. *MAT Exam Expository Papers* **2006**, 26.
- [27] Ma, J.; Dushoff, J.; Bolker, B. M.; Earn, D. J. Estimating initial epidemic growth rates. *Bulletin of mathematical biology* **2014**, *76*, 245–260.
- [28] Logistic function. 2021; [https://en.wikipedia.org/wiki/Logistic\\_function#cite\\_note-FOOTNOTECramer20023\T1\textendash5-2](https://en.wikipedia.org/wiki/Logistic_function#cite_note-FOOTNOTECramer20023\T1\textendash5-2).
- [29] Cramer, J. S. The origins of logistic regression. **2002**,
- [30] McKendrick, A.; Pai, M. K. XLV.—The rate of multiplication of micro-organisms: a mathematical study. *Proceedings of the Royal Society of Edinburgh* **1912**, *31*, 649–655.
- [31] Postnikov, E. B. Estimation of COVID-19 dynamics “on a back-of-envelope”: Does the simplest SIR model provide quantitative parameters and predictions? *Chaos, Solitons & Fractals* **2020**, *135*, 109841.
- [32] Reiser, P. A. Modified SIR Model Yielding a Logistic Solution. *arXiv preprint arXiv:2006.01550* **2020**,
- [33] Shigemoto, K.; Saito, T. A Logistic Curve in the SIR Model and Its Application to Deaths by COVID-19 in Japan. *medRxiv* **2020**,
- [34] Pelinovsky, E.; Kurkin, A.; Kurkina, O.; Kokoulina, M.; Epifanova, A. Logistic equation and COVID-19. *Chaos, Solitons & Fractals* **2020**, *140*, 110241.
- [35] Coronavirus en Ecuador: El país DUPLICA el Número de casos Y ya SON más De 22.000. <https://www.bbc.com/mundo/noticias-america-latina-52406283>.



- 
- [36] Coronavirus en AMÉRICA Latina: 7 gráficos para entender el AVANCE de la pandemia de covid-19 en la región. <https://www.bbc.com/mundo/noticias-america-latina-52405371>.
- [37] Wilson, W. G.; Lundberg, P. Non-neutral community dynamics: empirical predictions for ecosystem function and diversity from linearized consumer–resource interactions. *Oikos* **2006**, *114*, 71–83.
- [38] Gibson, W. T.; Wilson, W. G. Individual-based chaos: Extensions of the discrete logistic model. *Journal of theoretical biology* **2013**, *339*, 84–92.
- [39] Fogel, B.; Fogel, B. Ecuador: Coronavirus epicenter moves from Guayaquil to Quito. 2020; <https://brazilian.report/liveblog/coronavirus/2020/07/26/ecuador-coronavirus-epicenter-moves-from-guayaquil-to-quito/>.
- [40] León, J. M. Ecuador’s death toll during outbreak is among the worst in the world. 2020; <https://www.nytimes.com/2020/04/23/world/americas/ecuador-deaths-coronavirus.html>.
- [41] Millán Valencia, A. Coronavirus: ¿Por qué ecuador tiene el mayor número de contagios y muertos per cápita de covid-19 en sudamérica? <https://www.bbc.com/mundo/noticias-america-latina-52036460>.