Theme of the Thesis	Search for Invariant features on virus spread mathematical models: The case of Gompertz model and its renormalization
Abstract of Thesis Proposal	To study the evolution of a natural phenomenon we often use mathematical tools that are organized in several areas of knowledge, but which together constitute what we know as Mathematical Modeling. Thus, we are able to extract properties that allow understanding, enabling their interpretation, in order to satisfy a real need and obtain solutions to the various problems that result from the phenomenon. With the emergence of the global pandemic caused by the SARS-COV 2 virus, many were the models used by science to try to understand the evolution of the spread of the virus in order to help the health systems of each country to manage their resources efficiently, creating means to protect its population. One of these models was supported by the solution of systems of specific differential equations that became known as the Gompertz Model. The objective of this thesis is to organize the information related to the Gompertz model over time until it becomes the powerful tool that it currently represents in the study of viral infections, presenting a variation of the same model through the use of the Gamma function, creating a new model to be proposed to the scientific community. The new model presented has characteristics that originate in continuous dynamical systems, supported by differential equations, but also in stochastic and probabilistic systems, as well as models based on time series. The main objective of this thesis is to investigate the dynamics of a virus evolution created system, understanding the processes and the usefulness of it, as for example, the Gompertz model, with applications to non-continuous territories such as the Azores Islands.
Keywords	Differential Equations; Mathematical Modeling; Gompertz model; Gamma Function; Time Series;
Aims and Goals	The main objective of the thesis is to investigate the dynamics of mathematical models built based on Gompertz model system types, using as an initial step the Gompertz-Cabral model, used in viral prediction at the time of a pandemic in Portugal. It is intended to: adapt the modeling techniques, as much as possible, to a reality close to Azorean Portuguese society; draw comparative conclusions about viral evolution in various countries, using public data provided by the World Health Organization; Conclude on the structural changes needed to modify the proposed model, and similar models, in adapting studies to the various target audiences; Investigate the complexity of the model obtained, demonstrating through appropriate simulations, its impact and its efficiency in relation to the information already publicly known about the global pandemic in countries such as Portugal and others; Formulate all the basic mathematical structure that can give rise to a new model of viral evolution in a discontinuous territory, in general. The central research hypotheses are: (1) Is it possible to identify in the pandemic phenomenon, caused by Covid19, some mathematical Invariant features, which can be used to formulate an efficient model, allowing each decision-making country to save time and resources in the fight against its effects? (2) Is the Gompertz-Cabral model an efficient model to study the evolution of a virus in non-continuous territories such as the Azores Islands? It is intended, with the achievement of these objectives, in the development of all the methodological work of the thesis, in order to satisfy the two research hypotheses, to infer and obtain relevant information that is useful to decision-makers, in the decision-making processes, to optimize the protection of populations and management of the necessary resources in combating situations resulting from the spread of the disease caused by contagion.

Topic, Scope and Literature Summary

The impact of the Covid-19 pandemic on human health, social and economic life has aroused considerable interest in numerous scientific areas, including medicine, economics, psychology, mathematics and statistics, etc. Since the beginning of the pandemic caused by the Sars-Cov2 virus, and later by its variants, many studies have tried to predict the peak time and possible end time of new daily cases for different countries. These studies can be divided into two parts. The first part focused on finding an effective strategy to prevent or prevent or control the spread of the Covid-19 pandemic and how to reduce its harmful effects in all countries. Currently, the other part focuses on the vaccination process and the effects of the Covid-19 pandemic.

The outbreak of unusual viral pneumonia due to coronavirus was declared as a global pandemic by the World Health Organization (WHO), which emerged in the city of Wuhan, China, in late 2019. The Coronavirus, also known as Covid-19, has spread widely. rapidly around the world, despite all prevention efforts.

After the appearance of the first case, authorities in all countries quickly took some emergency measures to prevent society from its harmful effects and control its spread.

Despite all the measures and restrictions, the number of infected people in the world exceeded 268 million people, and the death toll more than 5 million people.

However, these precautions do not appear to be exactly successful or effective in controlling the disease so far. So while the global urgency has become to find a vaccine [1], the WHO has also used this crisis to identify vulnerabilities, modeling reported data from the Covid-19 outbreak around the world to predict and prevent similar events.

During this humanitarian crisis, humanity once again reminded that collecting epidemiological data and predicting epidemic trends are important for the development and measurement of public intervention strategies. The observations were aimed at estimating the scale and severity of Covid-19 and produced some epidemic prediction results by various mathematical models and prediction approaches to predict the transmission of Covid-19.

To address common questions such as exact predictions of infection growth rate, tipping point, duration of outbreaks, and total number of Covid-19 cases [2], researchers have proposed different mathematical approaches, including models based on the logistic function. simple [3], improved SIR and SEIR models [4], hierarchical polynomial regression models [5], statistical models including machine learning techniques [6], time-variable Markov process [7], Gaussian mixture model [8], among others.

One of the first studies in which mathematical methods were used to explain epidemiological dynamics is the study carried out by Kermack and McKendrick [9] in 1927. The authors developed a susceptible-infected-recovered equation model, known as the SIR model, to simulate the transmission of an infectious disease. This method consists of three ordinary differential equations (ODEs). Such mathematical models used in biological sciences can be summarized in [10-12]. In the SIR model, it is assumed that the population size is fixed and constant throughout the epidemic, also parameters such as age, sex, location and social behavior have no effect on the epidemic.

Let S = S(t) be the susceptibles those who can acquire the disease, but presently do not have it, I = I(t) be the infectives those who have the disease and can transmit it to the susceptibles, and R = R(t) be the removed class those who had the disease, are now recovered at time t. The SIR model can be described as

$$\frac{dS}{dt} = -\beta \frac{SI}{N},$$

$$\frac{dI}{dt} = \beta \frac{SI}{N} - \gamma I,$$

$$\frac{dR}{dt} = \gamma I,$$

where β is the transmission rate of the infection, γ is the rate of recovery and N is the total number of population.

An extension model of SIR model is the SIRD model first presented by Kermack and McKendrick [9], which also does not consider exposed, super-spreader, and asymptomatic populations but also has another compartment, namely Deaths (D). SIRD model can be explained detailed as

$$\begin{aligned}\frac{dS}{dt} &= -\beta \frac{SI}{N},\\ \frac{dI}{dt} &= \beta \frac{SI}{N} - \gamma I - \delta I,\\ \frac{dR}{dt} &= \gamma I,\\ \frac{dD}{dt} &= \delta I,\end{aligned}$$

where, additionally, δ is the death rate.

As stated above, the SIR model is not capable of capturing the incubation period, i.e., the time elapsed before developing symptoms. This leads to SEIR model which considers a transition from susceptible to exposed individuals [13].

Another model is SEIR model that is similar to the SIR model except that variable E is added for the fraction of individuals that have been infected but are asymptomatic, called as exposed individuals. This model can be described as follows:

$$\begin{aligned}\frac{dS}{dt} &= -\beta \frac{SI}{N},\\ \frac{dE}{dt} &= \beta \frac{SI}{N} - \alpha E\\ \frac{dI}{dt} &= \alpha E - \gamma I,\\ \frac{dR}{dt} &= \gamma I,\end{aligned}$$

α is the incubation rate (transition rate from E to I) [14].

Then in 2020, Zhang et al. [15] proposed their new model as "SEIQR" since they developed five compartments (susceptible-exposed-infectious-quarantine-recovered) to simulate the models which are including interventions such as social distancing and lockdown. In this model Q represents the number of quarantined people after the confirmation of the positive test result.

Besides all these epidemic models, some researchers focused on simple mathematical approaches based on the logistic equation, Gompertz equation, Bertalanffy equation, Richards model, Ratkowsky model etc. to study Covid-19 outbreak in different countries.

The Gompertz equation is given by

$$y = ae^{-be^{-ct}}$$

and the logistic equation is given by

$$y = \frac{a}{1+ce^{-bt}},$$

where a, b, c are the parameters of equations, t is time [16].

Several predictive models such as Gompertz, Logistic and Bertalanffy were applied to study Covid-19 in China by Jia et al. [17]. Using exponential law models, Zhang et al. [18] focused on the duration and onset of Covid-19 in several countries in Europe and America.

Pelinovsky et al. [19] compared two models based on logistic and Gompertz equations in describing the Covid-19 outbreak in 23 countries: Australia, Austria, Belgium, Brazil, Great Britain, Germany, Denmark, Ireland, Spain, Italy, Canada, China , Netherlands, Norway, Serbia, Turkey, France, Czech Republic, Switzerland, South Korea, USA, Mexico and Japan, using official data from the first wave of the coronavirus pandemic.

Furthermore, Wang et al. [20] predicted Covid-19 epidemic trends employing the logistic model and machine learning time series prediction model for Brazil, Russia, India, Peru and Indonesia. Aviv-Sharon et al. [21] worked on a generalized logistical model to account for the onset of Covid-19 in Asia.

Wu and others. [22] used the logistical growth model, the generalized logistical growth model, the generalized growth model and the generalized Richards model to make predictions of the Covid-19 outbreak in 29 provinces in China and the rest of the world. They also applied a logistic model to predict the peak of confirmed cases in Europe and the United States.
In addition, Villalobos [23] made some predictions for Costa Rica based on the goodness of fit of the logistic model and the Gompertz model for the data from China and South Korea.
Mendiete et al. [24] provided Covid-19 predictions for Cuba using logistic regression and Gompertz curves. The aim of this study is to adjust the logistic and Gompertz models to the distribution of Covid-19 in Cuba for confirmed and deceased cases, to explain the compatibility of these models.

Even so, with so many developments in this aspect of the development of mathematical models, there are still many open questions about the models themselves, and so far it has not been possible to standardize them. Therefore, a tool that allows this standardization will be the detection of factors that remain invariant in the various mathematical structures used by the various models, known as mathematical invariants in the field of research on nonlinear dynamical systems, as this would allow the creation of a base, a central nucleus of a model, thus being able to analyze the adaptive characteristics that would become more efficient.

One of the models that is very similar in the research of these invariants is, without a shadow of a doubt, the Gompertz-Cabral model, due to its mathematical structure and therefore it will serve as a starting point both for the research of the invariants, as well as to show the feasibility application of the models in countries such as Turkey. The Gompertz-Cabral model is a model that uses Gaussian functions and time series in its structure and relies on the help of all the information already collected by the World Health Organization to produce simulations that are easily comparable with real data.

Original Value, Importance of Thesis	In the vast scientific field of continuous dynamical systems, the search for mathematical invariants, which can serve as a basis for the construction of more efficient mathematical models, is a subject that has been developed with great intensity since the end of the 20th century and during these first years. decades of the 21st century. As is to be expected, the pandemic caused by COVID-19 was the main trigger that made it possible to accelerate the creation of technologies that could be used to understand the evolution of the spread of various viruses, but it was not the only one and so this issue continues to be today. very pertinent. In this search, the Gompertz model has always been in the shadows, not least because it was a model most used in the mathematical modeling of phenomena resulting from agricultural pests as well as in the management of maritime resources. However, throughout the pandemic evolution caused by COVID-19, it was possible to verify that this model had many characteristics that made it possible to create evolution scenarios for this virus. But none of the existing works so far, known and published in the usual scientific circles, sought to identify the mathematical invariant structures that endowed the same with this ability to build these scenarios, perhaps due to the complexity of the mathematical formulation, or perhaps due to the excessive use of tools. Thus, with the correct identification of these invariants and also with the help of the mathematical structure present in the Gompertz-Cabral model, it will be possible to efficiently modify the Gompertz model so that it can be adapted to the Azorean reality, or the Portuguese, in general, and not only a much more efficient way than is currently possible.

Methodology	As a base methodology, an analysis of the main mathematical models, currently described in the references, will be undertaken, simulating their execution with data from the World Health Organization, seeking to identify, with the help of graphics, temporal periods where it may be possible to identify certain patterns of common evolution. to the methods. These patterns will serve to identify the analytical mathematical components responsible for their emergence, organizing them by their proximity to the Gompertz model. There will then be an exhaustive analysis of all these components identified and their importance in each of the methods in which they are used. Mathematically, we will try to show the linear independence of the analytical tools, enabling the creation of a structural base that will constitute the candidate structure for the mathematical invariant of the original model to be created. The assertiveness of this candidate invariant will be tested by comparison with the Gompertz-Cabral model, using computer simulations with analytical support. With the support of other mathematical tools, such as Gaussian functions and Time Series, the invariant will be tested, using the pandemic evolution data in several countries around the planet. As each linear combination of the tools will generate a possible invariant, they will be submitted to a validation, accepting all those that meet the criteria of a statistical test of hypotheses with a degree of certainty greater than 95%. We will try to demonstrate that this group of tools will be an invariant to all similar mathematical models. The new model will then be developed, a variation of Gompertz's, and the various hypotheses of application to the study of pandemic evolution in Portugal, will be studied, showing its impact and efficiency by comparing it with other existing studies.
Widespread Impact - Added Value	Creation of a structural information database, containing possible mathematical invariants, recognizable patterns”, that increase the efficiency of the methods already known and others that can be created. Publication of scientific papers in international journals of mathematical modeling. Participation in international conferences focused on mathematical modelling.

Essential References

[1] Santos, D. and Gouva, W. (2021). Impact of virus genetic variability and host immunity for the success of Covid-19 vaccines. Biomedicine and Pharmacotherapy, 136, 111-272.

[2] Joby, M. and Mahanthesh, B. (2020). Logistic growth and SIR modelling of coronavirus disease (Covid-19) outbreak in India: models based on real-time data. Mathematical Modelling of Engineering Problems, 7(3), 345-350.

[3] Claudia, F. and Mortarino, C. (2021). The effect of swabs on modeling the first wave of the Covid-19 pandemic in Italy. Emerging Science Journal, 5, 37-61.

[4] Shaobo, H.; Peng, Y. and Sun, K. (2020). SEIR modeling of the Covid-19 and its dynamics. Nonlinear Dynamics, 101(3), 1667-1680.

[5] Ekum, M. and Adeyinka, O. (2020). Application of hierarchical polynomial regression models to predict transmission of Covid-19 at global level. International Journal of Clinical Biostatistics and Biometrics, 6 (1). doi:10.23937/2469-5831/1510027.

[6] Samuel, L.; Hussain, J. and Chhakchhuak, L. (2020). Applications of machine learning and artificial intelligence for Covid-19 (SARS-CoV-2) pandemic: a review. Chaos, Solitons & Fractals, 139, 110059.

[7] Joseph R., L.; Martin, M.R.; Martin, J.D.; Kuhn, P. and Hicks, J.B. (2020). Modeling the onset of symptoms of Covid-19. Frontiers in Public Health, 8. doi:10.3389/fpubh.2020.00473.

[8] Amit, S.; Singh, P.; Lall, B. and Joshi, S.D. (2020). Modeling and prediction of Covid-19 pandemic using Gaussian mixture model. Chaos, Solitons & Fractals, 138, 110023.

[9] Kermack, W.O.; McKendrick, A.G. (1927). A contribution to the mathematical theory of epidemics. Proceedings of the Royal Society of London, 1927, 115, 700-721.

[10] Brauer, F. and Castillo-Chavez, C. (2012). Mathematical Models in Population Biology and Epidemiology, Springer-Verlag, New York.

[11] Hethcote, H.W. (2000). The mathematics of infectious diseases. Society for Industrial Applied Mathematics Review, 42, 599--653.

[12] Murray, J.D. (2002). Mathematical Biology: I. An Introduction, 3rd ed.; Springer-Verlag, New York, 2002.

[13] Chowell, G. (2017). Fitting dynamic models to epidemic outbreaks with quantified uncertainty: a primer for parameter uncertainty, identifiability and forecasts. Infectious Disease Modelling, 2, 3, 379-398.

[14] Peng L., Yang W., Zhang, D., Zhuge, C. and Hong, L., (2020). Epidemic analysis of Covid-19 in China by dynamical modeling. arXiv:2002.06563

[15] Zhang, Y.; Jiang, B.; Yuan, J. and Tao, Y. (2020). The impact of social distancing and epicenter lockdown on the Covid-19 epidemic in mainland China: a data-driven SEIQR model study. medRxiv.

[16] Zeide, B. (1993). Analysis of growth equations. Forest Science, 39(3), 594--616.

[17] Jia, L.; Li, K.; Jiang, Y.; Guo, X.; Zhao, T. (2020). Prediction and analysis of coronavirus disease. arXiv:2003.05447v2, arXiv 2020.

[18] Zhang, X.; Renjun M.; Wang, L. (2020). Predicting turning point, duration and attack rate of Covid-19 outbreaks in major western countries. Chaos, Solitons & Fractals, 135.

[19] Pelinovsky, E.; Kokoulina, M.; Epifanova, A.; Kurkin, A.; Kurkina, O.; Tang, M.; Macau, E. and Kirillin, M. (2021). Gompertz model in Covid-19 spreading simulation. Chaos, Solitons and Fractals. doi: https://doi.org/10.1016/j.chaos.2021.111699

[20] Peipei, W.; Zheng, X.; Li, J. and Zhu. B. (2020). Prediction of epidemic trends in Covid-19 with logistic model and machine learning technics. Chaos, Solitons & Fractals, 139, 110058.

[21] Elinor, A.S.; Aharoni, A. (2020). Generalized logistic growth modeling of the Covid-19 pandemic in Asia. Infectious Disease Modelling, 5, 502--509.

[22] Wu, K.; Darcet, D.; Wang, Q.; Sornette, D. (2020). Generalized logistic growth modeling of the Covid-19 outbreak in 29 provinces in China and in the rest of the world. ArXiv200305681 Phys Q-Bio Stat., available at: http://arxiv.org/abs/2003.05681

[23] Jia, L.; Li, K.; Jiang, Y. Guo, X. (2020). Prediction and analysis of coronavirus disease. ArXiv200305447 Q-BioPE (under review; accepted; in press). available at: https:// arxiv.org/abs/2003.05447

[24] Medina, J.F.; Mendieta, M.S.; Corts, M.; Corts PhD.; Corts, M. and Iglesias M.S. (2020). Covid-19 forecasts for Cuba using logistic regression and Gompertz curves. MEDICC Review, 22(3).