Theory and Parameterization of Infections and Waves by Covid-19: A 6-Countries Data Analysis

Huber Nieto-Chaupis
Universidad Autónoma del Perú
Panamericana Sur Km. 16.3, Villa el Salvador, Lima Perú
Email: hubernietochaupis@gmail.com

2021 IEEE 21st International Conference on Bioinformatics and Bioengineering (BIBE) | 978-1-6654-4261-9/21/$31.00 ©2021 IEEE | DOI: 10.1109/BIBE52308.2021.9635536

***Abstract*—From data of USA, Japan, Germany, UK, Italy and Russian, it is claimed that the Global pandemic dictated by the dynamics of Corona virus exhibits distributions that would correspond to a morphology of Bessel-like type. Under the assumption that the pandemic contains phases of infection denoted by the velocity and acceleration of propagation of virus, then a model of polynomials given by the integer-order Bessel functions is proposed. These polynomials enter in a statistical approach to define the law of infections as function of time for the ongoing global pandemic. From this, the data evolution and their different behaviors are interpreted in terms of the different phases including the Delta variant for the recent months until August 2021.**

***Index Terms*—Covid-19, Bessel, Delta variant.**

I. Introduction

As it is well-known Corona Virus Disease 2019 (Covid-19 in short) has debuted in Wuhan China at the end of 2019 [1]. Subsequently, Covid-19 did its apparition in Milan (Italy) [2] and Spain. The pandemic becomes global after the confirmation of infections in thousands was done in USA, UK, Germany and Brazil. In according to official data, it was observed in some countries the peak of first wave along the months of April, May and June [3]. On the other hand it was also seen that at the months of July, August, September and October the number of infections was drastically reduced to a small fraction. To some extent the moderated number of infections can be seen as a tangible result from the social restrictions that was done in most countries since the confirmed knowledge that the disease is potentially transmissible through human-human interaction. It is because the exhaled aerosols of infected ones. At the beginning of 2021, an unprecedented case of massive infection was reported in India. In the end of first quarter of 2021, the so-called third wave came strongly in Germany, Russian and Japan, however it was seen strongly in India [4]. In the end of first semester, the third wave seems to be emerging in USA and South-America countries. In these countries it has been confirmed that that this new wave is entirely dictated by the Delta-variant. This mutation of Covid-19 seems to be rather infectious than previous strains of 2020 season as experienced by India. Because these facts, emerges the question: What is the pattern behind the global dynamics of Covid-19 in world? If any [5], then it is possible to visualize in it the main characteristics of ongoing pandemic? In this paper, from a theory based in statistical principles and the dynamical evolution of infections, a model of infections has been derived.

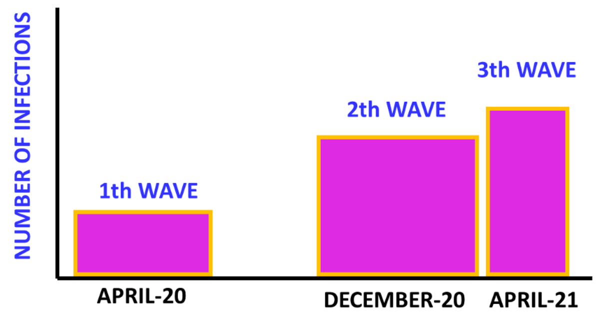

Fig. 1. Sketch of main pattern observed in data of most countries exhibiting a three waves pandemic.

From the fact that infections in time is the rate given by:

$$\mathbf{r}(\Delta t) = \frac{N_F - N_I}{\Delta t} \tag{1}$$

as commonly one see in country data, in the which the number of infections perceived as a difference $N_F - N_I = \Delta N$ is presented per day. Therefore one can write down:

$$\mathbf{r}(t) = \frac{N_F - N_I}{\Delta t} = \frac{\Delta N}{\Delta t} = \frac{dN(t)}{dt}. \tag{2}$$

From the fact that $\mathbf{r}(t)$ is continuous the one can apply subsequently additional derivatives that would be interpreted as the high-order changes of number of infections either by internal or external factors [6]. Once it has established the set of derivatives one can project it onto a framework of differential equations that would describe the time evolution of infections. Certainly this might to require to implement realistic initial conditions. Instead of this, in this study particular attention is paid onto the free parameters that would have direct implications on the matching between model and data. Actually the present investigation targets to propose a model in the which its free parameters play a critic role in the interpretation of theoretical distributions and its confrontation to data. The paper is structured as follows: In second section the theoretical machinery is presented. In third section the validation of model with official distributions of infections of some countries is done. In fourth section the analysis and interpretation of results is presented. Finally the conclusion of paper is given.

II. Theory of Global Pandemic

From Eq.2 if $\mathbf{r}(t)$ is a continue function at time, so that is a kind of velocity at the number of infections then it also might to admit its second derivative so that this can be seen as a kind of acceleration of number of infections. Mathematically speaking it is feasible to introduce an universal operator that satisfies a kind of eigenvalues equation:

$$\Lambda(m,t)\mathbf{n} = h(m,t)\mathbf{n} \tag{3}$$

In addition, at the representation $m, \Delta t$ the eigenvalue operation can be written as:

$$\Lambda(m,\Delta t) = \sum_{m=1} \Delta t^m \frac{d^m}{dt^2} \tag{4}$$

Now in the scenario that:

$$\Lambda(m > 2, \Delta t)\mathbf{n} = 0 \tag{5}$$

Then from Eq.4 and Eq.5 one arrives to:

$$\Lambda(m,\Delta t)\mathbf{n} = \left[\Delta t \frac{d}{dt} + \Delta t^2 \frac{d^2}{dt^2}\right]\mathbf{n} \tag{6}$$

thus for the approximation $\Delta t \rightarrow t$

$$\Lambda(m,t)\mathbf{n} = h(m,t)\mathbf{n} \tag{7}$$

so that one gets:

$$\left[\Delta t \frac{d}{dt} + \Delta t^2 \frac{d^2}{dt^2}\right]\mathbf{n} \Rightarrow \left[t \frac{d}{dt} + t^2 \frac{d^2}{dt^2}\right]\mathbf{n} = h(t,m)\mathbf{n}(t). \tag{8}$$

Clearly one can appeal to an infinite cases for $h(t,m)$ so that $\mathbf{n}$ might not be defined and its interpretation in praxis would lack of determinism and accuracy to the real data. Thus one can to take a "*adhoc*" decision and thus to opt by the function $\mathbf{h}(t,m)$=$m^2 - t^2$, fact that leads to write down the well-known integer-order Bessel equation [7]:

$$t^2 \frac{d^2\mathbf{n}(t)}{dt^2} + t\frac{d\mathbf{n}(t)}{dt} + (t^2 - m^2)\mathbf{n}(t) = 0. \tag{9}$$

This yields direct solution to $\mathbf{n}$:

$$\mathbf{n}(t) = J_m(t). \tag{10}$$

From this, one can illustrate to this point of study, the cases of velocity and acceleration of infections under the approach of integer-order Bessel functions,

$$\frac{d\mathbf{n}(t)}{dt} = \frac{1}{2}\left[J_{n-1}(t) - J_{n+1}(t)\right]. \tag{11}$$

In conjunction to this, in Fig.1 up and down panels are displayed up to the third derivative (up panel) and the square of 1th, 2th and 3th derivative (down panel).

It is noteworthy that this construction was possible with the assumption that $\mathbf{n}$ is free of uncertainties and biases, fact that that cannot acceptable. Therefore one should take into account the following:

- Because $\mathbf{n}$ would have to be associate to real data, the it would have an inherent statistical error due to the measurement of data.

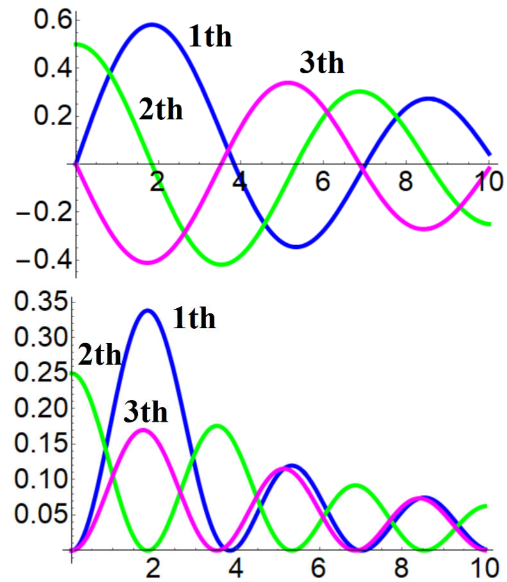

Fig. 2. Up panel, plot of 1th, 2th and 3th derivative. Down panel: the square of all three cases as shown above.

- The implementation of errors would also to include free model parameters that can improve or decrease the quality of model.

A. Linearity and Parametrization of Model

As mentioned above, $\mathbf{n}(t)$ in praxis are data that are updated daily in according to the evolution of pandemic. The register and counting of new cases cannot be absolute in the sense that the measurement of positivity is totally accurate. From the angle of the observation and measurement it is feasible to associate a statistical error that would have to decrease in time. Of course one would expect that a huge accumulation of data while pandemic is under time evolution, then the rate:

$$\frac{\mathcal{N}_{\text{WRONG}}}{\mathcal{N}_{\text{GOOD}}} << 1, \tag{12}$$

with $\mathcal{N}_{\text{WRONG}}$ the number of wrong reading (or ouput) of PCR, Antigen, Antibody, and RDT, the more popular fast Covid-19 tests as well as misinterpretations of test result. To propose a closed-form mathematical approach it is demanded a linear scenario. Therefore, with the implementation of free parameters a tentative proposal of statistical error on the measured number of infections can be written below as:

$$\frac{\sqrt{\mathbf{n}(t)}}{\mathbf{n}(t)} = \lambda(\beta - t) = \lambda\beta\left(1 - \frac{t}{\beta}\right) \tag{13}$$

exhibiting the morphology of a negative linear function. The mathematical proposal based on it emerges as purely logic because just in the so-called large pandemics while the systems are under time evolution, then one would expect a huge amount of data.Therefore the choice of form as Eq.13 is actually a desired scenario that encompasses in acoording to one might to experince in large pandemics: the statistical errors would have to be small in time. From this plus Eq.10 one solves for $\sqrt{\mathbf{n}(t)}$ yielding:

$$\sqrt{\mathbf{n}(t)} = \lambda(\beta - t)\mathbf{n}(t) = \lambda(\beta - t)J_n(t). \tag{14}$$

When $\sqrt{\mathbf{n}(t)}$ denotes the error coming from systematics and misinterpretation the one can redefine it by $\sqrt{\mathbf{N}(t)}$ under the assumption that the maximum of $\mathbf{n}(t)$ is depending on an extra parameter or "width of wave" Δ, so that one arrives to:

$$\sqrt{\mathbf{N}(t)} = \lambda(\beta - t)\mathbf{n}(t) = \lambda(\beta - t)J_n\left(\frac{t}{\Delta t}\right)$$
$$\Rightarrow \mathbf{N}(t) = \lambda^2(\beta - t)^2 J_n^2\left(\frac{t}{\Delta t}\right). \tag{15}$$

It should be noted that a compact form of Eq.15 can be written below as:

$$\mathbf{N}(t) = H^2(t, \beta, \lambda) J_n^2\left(\frac{t}{\Delta t}\right). \tag{16}$$

In order to illustrate Eq.15, in Fig.3 a toy model when the free parameters λ=β=Δt=1.0. The embedded panel is showing the oscillating behavior of Bessel function that for this exercise the integer $n = 1$ was taken. The average view of Fig.3 is clearly showing an growth with respect to time. In this manner, the entire dynamics of infection might be contained in the free parameters.

B. Role of Free Parameters

In order to project Eq.15 onto official data of Covid-19 pandemic, it is needed to associate a firm role to each one of free parameters:

1) λ: It is strongly linked to the amplitude of full distribution. Thus, while it is small then the peaks of all waves would have also to be shorter, and in the contrary. The values that can take are inside an acceptable scenario.

2) β: This is related to the asymmetry between a first wave and the subsequent ones that would appear in the following periods. Thus, for higher values of β the peak in the first wave shall be higher than the next ones. In the inverse case, the lowest values would produce a distribution with the highest talls of all peaks, as well as it is possible to expect the apparition of more waves.

3) Δt: It is perhaps the most important free parameter because its variation inside a short window would reproduce the ongoing dynamics a seen recently in the apparition of the so-called Delta-variant of Covid-19. In contrast to previous parameters, Δt moves in a small range given a main value. In Fig.4 the case when $\lambda = 0.075$, $\beta = 110$ and $\Delta t = 25$ is shown. Here are emphasized the role of free parameters. It is noteworthy the role of Δ that is linked to the apparition of third wave and presumably to the one of so-called Delta-variant. In Fig.5 same case but in a manner of contour plot is displayed. The peaks of a possible scenario of first and second waves are indicated by the arrows.

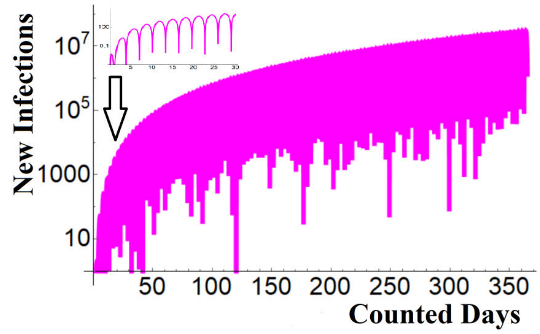

Fig. 3. Plotting of Eq.15 for the first order of Bessel function with $\lambda = \beta = \Delta t = 1.0$. The inner panel is a magnification for the first days.

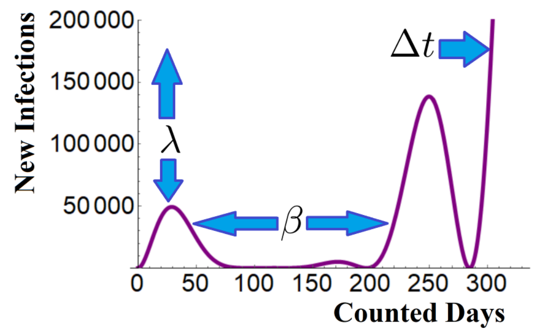

Fig. 4. A possible case of distribution of new infections versus counted days in the ongoing Covid-19 pandemic. Here all three free parameters are showing their role in the present theoretical model.

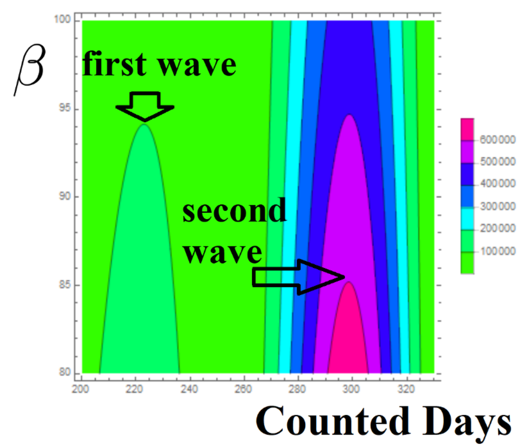

Fig. 5. Contour plot of Fig.4 with the arrows indicating the peaks of first and second waves of pandemic. This was done with the assistence of Wolfram [8].

TABLE I
FITTED VALUES FOR 6 COUNTRIES DATA.

Country	λ	β	Δt	waves	Error-δ	δ-variant
USA	0.030	10	25	3	12%	yes
UK	0.004	10	45	4	23%	yes
Germany	0.0075	80	18	3	25%	no
Russia	0.0095	110	55	3	28%	yes
Japan	0.0055	132	24	4	19%	yes
Italy	0.0085	80	25	2	36%	no

III. Model Validation

In Fig.6 the official distributions of Covid-19 cases for the USA [9] and UK [10] data are displayed. One can see the similarity between them despite the fact of spatial distance. Interestingly one can see that the peaks are just around the same week. Due to this, the modeling have yielded the values of the free parameters as given in Table-I. It is clear the large difference of the values of λ by one order of magnitude. Indeed both countries share same value of β revealing that have same asymmetry. it is approximately same number of cases in both first and subsequent waves. On the other hand, UK exhibits a suspected large peak on July first week, fact that can be seen as the beginning of an aggressive fourth wave due to the new "delta-variant" of ongoing Corona virus. Under the assumption that flights between London and East-Coast cities (New York and New Jersey, for example) have been running along June and July, then it is feasible to expect that the beginning of the step-shape in the USA along August data, can be strongly related to imported cases mainly from UK. In Fig.7 the data pair given by Germany [11] and Italy [12] are displayed. Interestingly as the previous case given by USA and UK, one can see that both Germany and Italy are sharing approximately same morphology in the distributions of new cases. Because this the parameter λ have a difference of 0.0010. Also $\beta = 80$ for both countries However, Italy exhibits a second wave with a large width in comparison to Germany. Furthermore, the beginning of a third wave emerges on the July 1th as seen in data. Nevertheless in contrast to USA and UK it is not evident of the beginning of a wave linked to the whole pandemic. Instead of this, one can relate this to the presence of a new strain with different infection capabilities. Thus one can argue that in these 2 countries Covid-19 pandemic has finished at the end of June. To claim that this is valid one can compare the beginning of second wave that seems to be similar to a step-function whereas the beginning at the end of June **contains a well-defined tail**. Thus, while not any evidence of a step-function at the beginning of July then one can talk about (in a hypothetically manner) as the beginning of a new pandemic. In Fig.8 one can see there is an evident strong difference between data of Russia [13] and Japan [14]. In fact, the value of λ for Russia is twice than the one for Japan. Although for these countries the value for β is hardly different each other, again the value of Δt for Russia is twice than the Japan case. Because this, Japan exhibits up to 4 well-defined waves corresponding to same strain (including the so-called "Delta-variant"). A fact that appears with a notable clarity is the why Russia and Japan are presenting same morphology in their distributions. One can answer this question starting from that either Moscow and Tokio are canceling their flights each other, or there is an decreasing at the number of flights between them. Interestingly one can seen the fast increasing of new infections at July, fact that strongly coincides to the period of Olympic Games that without any doubt has been the main cause of new outbreak of pandemic in Japan. Thus, this paper claims that "Delta-variant" could have entered to Japan in the different arrived planes very probably of all countries that already were exhibiting the outbreak of "Delta-variant" such as UK (it is taken as example in according to data presented in Fig.6, Fig.7 and Fig.8).

A. Estimation of Error

The calculation of error has been done through the usage of following equation:

$$\delta = \sqrt{\frac{\sum_q^Q \left[\left(\frac{N_T-N_D}{\Delta_T-\Delta_D}\right)^2\right]_q + \sum_m^M \left[\left(\frac{N_T-N_D}{\tau_T-\tau_D}\right)^2\right]_m + \sum_j^J \left[\left(\frac{\mathbf{R}_T-\mathbf{R}_D}{r_T-r_D}\right)^2\right]_j}{Q+M+J}} \tag{17}$$

With δ is estimated by country. Here N_T and N_D denote the number of infections at the observed Q waves for the theory and data, respectively. The widths Δ_T and Δ_D are the values of the width the peaked periods for M peaks, and τ_T and τ_D are the estimated days by which the peak is observed. To complete the error, the quantities $\mathbf{R}_T$ is an aleatory point belonging to the of theoretical distribution (E.15) that is compared to data to any day $\mathbf{R}_D$ inside the whole dataset that are taken in a random manner. For this exercise up to J =18 points at the ordinate and abscissa were considered. The criterion of validity of Eq.17 breaks when:

$$\sum_j^J \left[\left(\frac{\mathbf{R}_T-\mathbf{R}_D}{r_T-r_D}\right)^2\right]_j$$

$$> \sum_q^Q \left[\left(\frac{N_T-N_D}{\Delta_T-\Delta_D}\right)^2\right]_q + \sum_m^M \left[\left(\frac{N_T-N_D}{\tau_T-\tau_D}\right)^2\right]_m \tag{18}$$

by the which is indicating that the aleatory choice cannot be greater than evidence. In all those scenarios that are near to this issue one can to decrease the value of J in order to reach positive values for the square root. Nevertheless emerges another issue that would not allow to carry out this "blind" action at the sense that there is not any preference period with respect to the wave. In other words, Eq.18 tells that statistical errors belonging to a global pandemic is an issue of deterministic character and it cannot be stochastic.

IV. Conclusion

In this paper, a theory approach concerning to the apparition of waves in global pandemic has been presented. The modeling has considered as main arguments the fact that data of new cases of infections by Covid-19 are exhibiting a well-defined pattern that would be adjusted to an integer-order Bessel function. In this way, a mathematical model that fits to some countries data was presented. It was extracted the free parameters of model up to for 6 countries. Finally, the error of modeling was estimated. Due to the rapid increasing of new cases because the apparition of ”Delta-variant” as well as the Japanese data, this paper concludes that still the intercontinental flights is a strong cause to disperse the strain in the world.

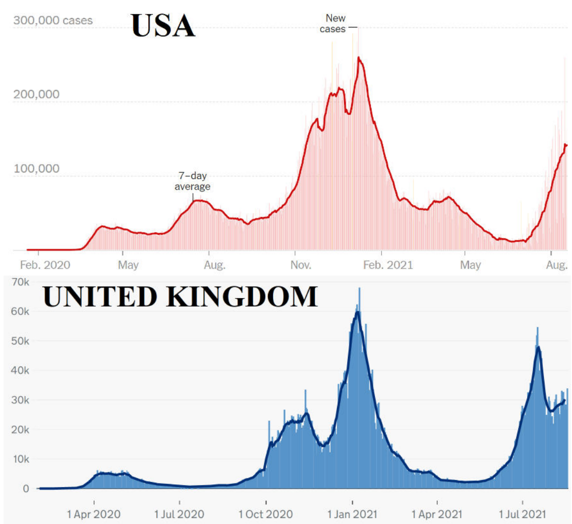

Fig. 6. Top: Data of infections by Covid-19 USA until 20s of August, by showing up to three well-defined waves. The rise of ”Delta-variant” is seen at the middle of August. Down: Data of infections by Covid-19 USA until beginning of August, by exhibiting 4 waves with the outbreak of ”Delta-variant” already running.

References

[1] Lin-Qi, et.al., Factors associated with the duration of viral shedding in adults with COVID-19 outside of Wuhan, China: a retrospective cohort study, Diseases Volume 96, July 2020, Pages 531-537.

[2] The response of Milan’s Emergency Medical System to the COVID-19 outbreak in Italy, The Lancet, Vol 395, Issue 10227, E49-E50, March 14, 2020.

[3] Huber Nieto-Chaupis, Identifying Second Wave and New Variants of Covid-19 from Shannon Entropy in Global Pandemic Data, 2021 Fifth World Conference on Smart Trends in Systems Security and Sustainability (WorldS4).

[4] Rajani Kumari, et.al., Analysis and predictions of spread, recovery, and death caused by COVID-19 in India, Big Data Mining and Analytics, 2021, Volume: 4, Issue: 2.

[5] Martijn J.Hoogeveen and Ellen K.Hoogeveen, Comparable seasonal pattern for COVID-19 and flu-like illnesses, One Health, Volume 13, December 2021, 100277.

[6] Xingyu Li,et., Mathematical model of the feedback between global supply chain disruption and COVID-19 dynamics, Scientific Reports volume 11, Article number: 15450 (2021).

[7] Watson, G. N, A Treatise on the Theory of Bessel Functions, Cambridge University Press; 2th Ed. (1 Agosto 1995).

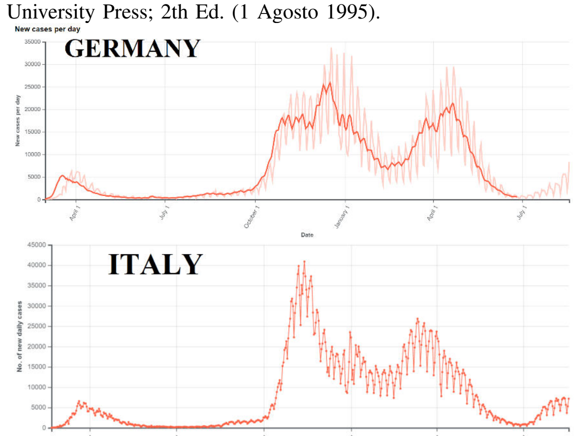

Fig. 7. Up: Number of infections by Covid-19 at Germany. Up to three peaks can be seen without any clue to claim that an outbreak by ”Delta-variant” has initialized. Down: The case of Italy that exhbits only two well-defined waves. The shape of distributions is rather similar to the case of Germany. Italy also might do not show no large data in a third wave dictated by the ”Delta-variant”.

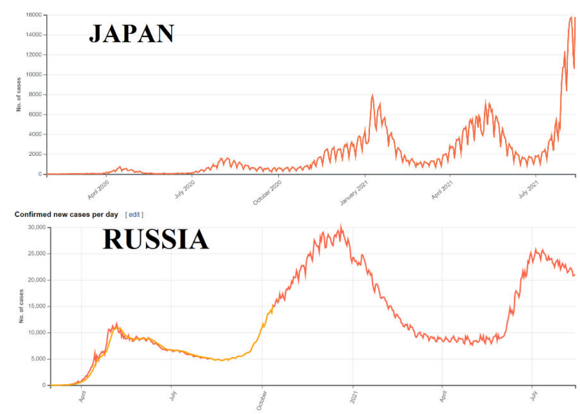

Fig. 8. Up: Data of Covid-19 infections from Japan. Here one can see up to three different waves with the last one presumably due to the ”Delta-variant”. Also one can see the asymmetry of waves. Down: Data of infections by Covid-19 from Russia. Here up to three waves can be seen. According to this data, no any track of the ”Delta-variant”.

[8] https://www.wolfram.com/mathematica/

[9] https://github.com/nytimes/covid-19-data

[10] https://coronavirus.data.gov.uk/details/cases

[11] https://www.zeit.de/wissen/corona-karte-deutschland-aktuelle-zahlen-landkreise

[12] https://www.epicentro.iss.it/coronavirus/bollettino/Bollettino-sorveglianza-integrata-COVID-19 18-agosto-2021.pdf

[13] https://graphics.reuters.com/world-coronavirus-tracker-and-maps/es/countries-and-territories/russia/

[14] https://covid19.who.int/region/wpro/country/jp