Mathematical Modeling and Estimation for Next Wave of COVID19 in Poland

Arti M.K.

NSUT: Netaji Subhas University of Technology

Antoni Wilinski (✉ awilinski@wsb.gda.pl)

WSB University in Gdańsk <https://orcid.org/0000-0002-2817-9095>

Research Article

Keywords: COVID19, Prediction, Mathematical Modelling, Gaussian Mixture Model.

Posted Date: August 19th, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-814363/v1>

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

[Read Full License](#)

Mathematical Modeling and Estimation for Next Wave of COVID19 in Poland

Arti M.K.¹ and Antoni Wilinski^{2*}

1 – NSUT East Campus (AIACTR), New Delhi, India

2 – WSB University in Gdansk, Poland, awilin@o2.pl

Abstract

We investigate the problem of mathematical modeling of new corona virus (Covid19) in Poland and tries to predict the upcoming wave. A Gaussian mixture model is proposed to characterize the COVID-19 disease and to predict a new / future wave of COVID-19. This prediction is very much needed to prepare for medical setup and continue with the upcoming program. Specifically, data related to the new confirmed cases of COVID-19 per day are considered, and then we attempt to predict the data and statistical activity. A close match between actual data and analytical data by using the Gaussian mixture model shows that it is a suitable model to present new cases of COVID-19. In addition, it is thought that there are N waves of COVID-19 and that information for each future wave is also present in current and previous waves as well. Using this concept, predictions of a future wave can be made.

Index Terms: COVID19, Prediction, Mathematical Modelling, Gaussian Mixture Model.

1. INTRODUCTION

Currently, the world is threatened by new upcoming of waves of Covid19, a new disease transmitted by the corona family virus. Many countries around the world have noticed a large number of cases of Covid 19 from December 2019 onwards Let us recall some basic facts about the pandemic and recommended behaviors. People with low immunity, aging, and lung-related medical problems are more prone to Covid 19 infections [1,2,3,4]. Symptoms of Covid19 are coughs, colds, respiratory problems that are very similar to the flu. It is clear from doctors that a person infected with Covid19 recovers within 14 -16 days because the incubation period of the novel corona virus is 14 days. Covid19 protective measures should be protected by frequent hand washing, to avoid touching the mouth, nose, and face, and to keep the public (2 m or 6 feet) away from other people. Covid19 is now a pandemic as announced by the World Health Organization (WHO) [1]. Therefore, preparation for health care should be adequate worldwide. As the disease is contagious, the number of young infected people is increasing rapidly. If hospitals are not well prepared, health workers will not be able to function properly. In this case, it is inevitable that we will have an accurate measure of new Covid19 cases, which can help medical managers and administrators. Now three waves of Covid 19 have appeared in Poland and in several other countries as well. An important question about the current state of Covid 19, that is, when will the next wave of Covid 19 emerge?

The conviction of the Polish authorities is rather optimistic, although very cautious. The daily announcements of the ministry of health (July 2021) predict the fourth wave of the pandemic, most often including fall dates. However, it results rather from premonitions and informal forecasts of an increase in prohibitions, e.g. in connection with the beginning of the school year in the

stationary mode planned by the authorities. The anticipated appearance of another wave of infections is also reinforced by the growing aversion to vaccinations of the so far unvaccinated part of the population still below the threshold of general social immunity (in July 2021, about 41% of the population was vaccinated with the second dose [34]). Therefore, virus spread models and prediction methods are still an extremely important challenge for the world of science, including the part that uses statistical inference and artificial intelligence. The authors of this work considered confirmed cases as the most important time series, ignoring others, which does not mean disregarding them, but only an attempt to model reality in this way - with one variable accumulating the influence of many predictors.

In previous publications, the problem of predicting the spread of the virus in Poland has been discussed many times [19, 24, 26, 32], often along with the analysis of the situation in other countries [14, 22, 23, 25]. In the above-mentioned studies, statistical data on Poland are directly used to exemplify changes in the spread of COVID-19 [19, 24, 26, 31, 33]. Moreover, there are many studies presenting general models that can be applied to quite any country and the selected data are only a pretext for confirming the methodological correctness of the applied model and the prognostic method [7, 12, 13, 16, 18, 27, 28, 29, 30, 31]. Some of them are devoted to the analysis of the situation in a selected country, eg in the USA [10, 13] or India [8, 9]. Interesting and inspiring may be the analysis in terms of the applied models and predictive methods, mainly emerging from artificial intelligence and machine learning algorithms.

Data for work were collected from sources such as [1, 5, 6, 34].

In this paper, we propose a Gaussian composite model to present new confirmed cases of Covid 19. Limited estimates can be made using this model. We have read and discussed details of Poland. The proposed model can also be used in other countries as well. Approval intensity is demonstrated by comparing available data with analytical results. In addition, an estimate of the next wave of COVID-19 wave can be made using this model.

2. PROPOSED MODEL

We consider that there are total N Gaussian waves. Let us write a mixture Gaussian function with the help of [21] as

$$X(t) = A_1 + A_2 t + \sum_{k=1}^N p_k e^{-\frac{(t-m_k)^2}{2sig_k^2}}, \quad (1)$$

where A_1 and A_2 are constants, e is the exponential function; p_k , m_k , and sig_k are the weight, mean, and standard deviation of k -th Gaussian wave. It can be noticed from (1) that $X(t)$ contain the information of all N Gaussian waves, i.e., all previous, present and future waves. Relation (1) is a general relation and by choosing different values of p_k , m_k , and sig_k , characterization of a particular country can be obtained. Let us consider the case of Poland, in this case, we can write (1) as

$$X(t) = 10 + t + 25000 e^{-\frac{(t-255)^2}{550}} + 12000 e^{-\frac{(t-305)^2}{350}} + 32000 e^{-\frac{(t-390)^2}{550}} \quad (2)$$

Fig. 1. Analytical and actual plot of daily new cases of Covid 19 in Poland.

Actual data of daily new cases of Covid 19 in Poland and analytical data is plotted in Fig. 1. Analytical data is calculated by using (2). The actual data of daily confirmed cases in Poland of Covid 19 is considered from 04/03/2020 to 09/07/2021 [5]. It can be noticed from actual data [5] that first peak of Covid 19 was observed on 07/11/2020 in Poland with 27875 fresh cases, whereas the analytical value is approx. 26570, which is quite close to the true value. Further, a spike can be seen in actual data of value 37596 on 23/11/2020, it is not modelled properly with the analytical relation given by (2). Then a very small second wave can be observed from Fig. (1), which started in mid of December 2020 and vanished in first week of February 2021. It's peak value was around 14,216, as seen from the figure. Third wave of Covid 19 started in the month of February 2021, just after second wave. It was more severe as compared to first and second wave. It's peak value was around 35,246 [5] and it was observed in last week of March 2021. If we calculate the analytical peak of third wave, it is approximately 33,000, which is very close to the actual value. Therefore, it can be concluded that (2) characterizes the new cases of Covid 19 in Poland efficiently.

3. RESULTS OD THE PRDICTION OF THE NEXT WAVE OF COVID-19

In this section, it is proposed to estimate the forthcoming Covid 19 wave, in particular, proposing a predictor of the next covid wave along with its maximum value. To explain this approach, let us first consider the measurement of mean, standard deviation, and the peak value of a Gaussian wave by utilizing its sample values.

3.1. Estimation of mean, standard deviation, and the peak value of a Gaussian wave:

We can write a Gaussian function as

$$Y(t) = p e^{-\frac{(t-m)^2}{2s^2}}, \quad (3)$$

where p , m , s are the height, mean, and standard deviation of the Gaussian wave provided by (3). The plot of (3) is also shown in Fig. 2. Let us assume that some samples of $Y(t)$ are available as shown in Fig. 3 and we need to estimate p , m , s by using these samples. From (3), we obtain

$$Y(t_1) = p e^{-\frac{(t_1-m)^2}{2s^2}}, \quad (4)$$

$$Y(t_2) = p e^{-\frac{(t_2-m)^2}{2s^2}}, \quad (5)$$

and

$$Y(t_3) = p e^{-\frac{(t_3-m)^2}{2s^2}}. \quad (6)$$

By using (4) and (5), we get

$$\log_e \frac{Y(t_1)}{Y(t_2)} = \frac{(t_1+t_2-2m)(t_2-t_1)}{2s^2}. \quad (7)$$

Fig. 2. Plot of Gaussian Wave

Fig. 3. Samples of Gaussian function given by (2)

Similarly, by using (5) and (6), we can write

$$\log_e \frac{Y(t_2)}{Y(t_3)} = \frac{(t_2+t_3-2m)(t_3-t_2)}{2s^2} . \quad (8)$$

With the help of (7) and (8), after some algebra, we obtain

$$m = \frac{(t_2^2-t_1^2) \log_e \frac{Y(t_2)}{Y(t_3)} - (t_3^2-t_2^2) \log_e \frac{Y(t_1)}{Y(t_2)}}{2((t_2-t_1) \log_e \frac{Y(t_2)}{Y(t_3)} - (t_3-t_2) \log_e \frac{Y(t_1)}{Y(t_2)})} \quad (9)$$

$$\text{and } s = \sqrt{\frac{(t_1+t_2-2m)(t_2-t_1)}{2 \log_e \frac{Y(t_1)}{Y(t_2)}}} . \quad (10)$$

By using (4), (5), and (6), we get

$$p = \frac{Y(t_1)+Y(t_2)+Y(t_3)}{e^{-\frac{(t_1-m)^2}{2s^2}} + e^{-\frac{(t_2-m)^2}{2s^2}} + e^{-\frac{(t_3-m)^2}{2s^2}}} . \quad (11)$$

The value of p can be obtained by substituting the values of m and s from (9) and (10), respectively, into (11). Further, it can be observed that only three samples are sufficient to get a Gaussian wave.

3.2. Estimation of mean, standard deviation, and the peak value of a Gaussian wave with N samples:

In this subsection, we assume that N samples of a Gaussian function is available, as in the case of Covid 19 new wave, in that case following procedure can be used:

- 1) Divide all samples in the group of three, if it is not possible leave some old samples and take all recent samples, which can make a group of three.
- 2) Estimate p, m, s by using the method given in subsection A for all groups.
- 3) Estimate p, m, s by performing averaging of p, m, s obtained by different groups.

3.3. Prediction of next wave of Covid 19:

In the current situation of Covid 19, the most important question is that when will next wave of Covid 19 start for the proper arrangements performed by hospitals and other authorities as well. It can be noticed from (1) that any in any sample of $X(t)$, i.e. from the information of fresh cases of Covid 19, one can have many more details of Covid 19. It is because of the fact that the $X(t)$ contains the data of preceding waves, present Covid 19 wave and upcoming Covid 19 waves also. Another important fact is that Gaussian function decreases as it moves away from its mean value, therefore, it can be assumed that the governing waves are only nearby Gaussian waves, all other will contribute negligibly. Let us rewrite (1) in more detailed form as

$$X(t) = A_1 + A_2 t + p_1 e^{-\frac{(t-m_1)^2}{2sig_1^2}} + p_2 e^{-\frac{(t-m_2)^2}{2sig_2^2}} + p_3 e^{-\frac{(t-m_3)^2}{2sig_3^2}} \dots + p_N e^{-\frac{(t-m_N)^2}{2sig_N^2}}, \quad (12)$$

Where N can take any positive integer value, large or small. It can be seen from (12) that when the first Covid 19 wave was detected, it contains some information for all future Covid 19 waves. But much of the information is about the second wave of Covid 19, that is, the coming wave. Details of the remaining waves such as third, fourth etc. was very small or overlooked due to the Gaussian nature of the Covid 19 waves. Similarly, when second Covid 19 wave started, these samples contain specific details of the third Covid 19 wave and so on. Using this concept, estimates of the next Covid 19 wave can be made. We begin the measurement process by assuming that some Covid 19 samples are available. The first wave has not yet fully arrived. From (12), we can write

$$X(t) = A_1 + A_2 t + p_1 e^{-\frac{(t-m_1)^2}{2sig_1^2}}, \quad (13)$$

In (13), A_1 and A_2 can be obtained by curve fitting.

It is assumed that first wave has not been fully arrived, so by following the method given in III-A, we can find the values of p_1, m_1 and sig_1 . As more samples of first Covid 19 wave received, estimation of p_1 will be better. Moreover, these values will also supply

some details of second Covid 19 wave also. If there is no upcoming Covid 19 wave, then $X(t)$ must be zero after the first Covid 19 wave. It shows that after sufficient time, i.e., after dying out the first wave, there must be zero or very small daily new cases of Covid 19 patients. But if it is not so then $X(t) - p_1 e^{-\frac{(t-m_1)^2}{2\sigma_1^2}}$ are the values which may generate second wave of Covid 19. By using these sample values, one can estimate the upcoming covid 19 wave and so on.

3.4. Prediction of next wave of Covid 19 in Poland:

We want to estimate about the next or fourth Covid 19 wave in Poland. From (2), we have the characteristics of the first, second, and third waves of Covid 19. The calculation of the parameters of the second wave and third wave of Covid 19 is performed using the method given in section III-A. The relationship of new Covid 19 cases daily is given in (2). If there is no fourth wave then third wave will die out completely and new covid cases count will be zero. But if it is not so then we need to calculate

$$X_1(t) = X(t) - 32000 e^{-\frac{(t-390)^2}{550}}. \quad (14)$$

By using the sample values of $X_1(t)$, new covid wave can be estimated as

$$X(t) = 10 + t + 25000 e^{-\frac{(t-255)^2}{550}} + 12000 e^{-\frac{(t-305)^2}{350}} + 32000 e^{-\frac{(t-390)^2}{550}} + 6540 e^{-\frac{(t-740)^2}{350}}. \quad (15)$$

Fig. 4. Estimated fourth Covid 19 wave in Poland by considering the new cases from 03/04/2021 to 18/05/2021.

Fig. 4 shows a plot of analytical and actual cases of daily new cases of Covid 19 in Poland by considering the data of new daily cases from 03/04/2021 to 18/05/2021. Further, it also shows the estimated fourth Covid 19 wave in Poland. It can be observed from Fig. 4 that next or upcoming wave of Covid will be comparatively less severe and it may be expected to appear in the month of October 2021 or November 2021. Moreover, it is very early stage and it is shown by simulations that as well as more samples have been received, a better estimation can be performed. The next Covid 19 wave will also depend upon several other factors like social distancing, use of mask etc. It can be concluded that the estimated value of mean, variance, and height of the Gaussian wave depend largely on the set of samples used.

Fig. 5. Estimated fourth Covid 19 wave in Poland by considering the new cases from 03/05/2021 to 07/06/2021.

Fig. 5 shows a plot of daily new cases of Covid 19 versus days in Poland by considering the data of new daily cases from 03/05/2021 to 07/06/2021. Further, it also shows the estimated fourth Covid 19 wave in Poland during the month of November or December 2021. As seen from Fig. 4 and 5, that the prediction of upcoming next wave is largely dependent upon the data samples used. Further, it can also be observed that almost one month delay or advancement can be noticed, if we change set of sample values.

4. DISCUSSION AND CONCLUSIONS

A mathematical model has been proposed to characterize the Covid 19 daily new cases. A close match between actual and analytical data is evident from numerical results provided in the paper. By choosing different parameter values in the proposed function, the plots can be obtained for other countries also. Specifically, the characterization has been performed of Covid 19 daily new cases in Poland. Moreover, an estimation method has been developed for Gaussian function by utilizing its samples. It can be concluded that the estimated mean, variance and peak of Gaussian function largely depends upon the set of samples used.

The results obtained for Poland do not seem to be in line with the so-called common expectation. However, the authors present and justify the method, obtaining results that are consistent with historical data so far. The model does not take into account many potentially significant predictors that would e.g. occur in regression models (e.g., time series of tests, deaths, recovered, vaccinations, restrictive actions, health service status, etc.). In this matter, it is worth returning to valuable already published and partially cited works [7, 9, 13, 16, 18, 27, 29].

Most of the works cited here, those aimed at forecasting methods, limit their aspirations to rather short prediction horizons, often several days, weeks, and less often months, usually stipulating a wide range of the possible forecasted result. In the model with the use of Gaussian waves, we are far ahead, perhaps risking our reputation. However, we show that such far predictive horizons can be achieved, as usual, with the risk of error. We also assume that there are countries for which this method will be more effective than for Poland, as well as those for which it will be less successful. Here is an example of the philosophy of computational intelligence.

With large data set, estimation will also be improved. In addition, the process of estimating the upcoming Gaussian wave has been also proposed. It is shown by estimation that with more samples obtained, a better estimate can be made.

DECLARATIONS:

Funding – not applicable;

Conflict of interest: The Authors declare no conflict interests.

Availability of data and material:

The data used in the work is available on the following pages:

<https://www.who.int/emergencies/diseases/novel-coronavirus-2019>

<https://www.worldometers.info/coronavirus/countries>

<https://ourworldindata.org/covid-vaccinations> 2019

www.pokazwirusa.pl/wykresy

Code availability – on request;

Author's contributions:

Arti MK.: Conceptualization, Methodology, Data Curation, Writing, Software;

Antoni Wilinski: Conceptualization, Validation, Writing, Review and Editing;

Ethics approval: The authors are familiarized with the ethical responsibilities of Authors and meet these conditions.

Consent to participate – yes;

Consent to publication – yes;

REFERENCES

- [1] <https://www.who.int/emergencies/diseases/novel-coronavirus-2019> [Accessed on May 12, 2021].
- [2] P. Zhou et al. A pneumonia outbreak associated with a new coronavirus of probable bat origin. *Nature* 579, 270–273 2020. <https://doi.org/10.1038/s41586-020-2012-7> Medline
- [3] F. Wu et al. A new coronavirus associated with human respiratory disease in China. *Nature* 579, 265–269 2020. <https://doi.org/10.1038/s41586-020-2008-3> Medline
- [4] A. E. Gorbalenya et al. Severe acute respiratory syndrome-related coronavirus: The species and its viruses – a statement of the Coronavirus Study Group. *Nat. Microbiol.* 2020. <https://doi.org/10.1038/s41564-020-0695-z>
- [5] <https://www.worldometers.info/coronavirus/countries> [Accessed on July 10, 2021].
- [6] <https://ourworldindata.org/covid-vaccinations> 2019 [Accessed on July 11, 2021].
- [7] Manav R. Bhatnagar, “COVID19: Mathematical Modeling and Predictions,” ResearchGate, DOI:10.13140/RG.2.2.29541.96488, published on April 02, 2020.
- [8] Arti, M. K. and Kushagra Bhatnagar, “Modeling and predictions for COVID 19 spread in India. <https://doi.org/10.13140/RG.2.2.11427.81444> (2020)
- [9] S. Roy and K. R. Bhattacharya, “Spread of COVID-19 in India: A Mathematical Model”, ResearchGate, DOI: 10.13140/RG.2.2.15878.52802, published on April, 2020.
- [10] Sinha, D. and Klahn, N, Mathematical Modeling Study of the 2020 CoVID-19 Outbreak in the United States, SSRN Electronic Journal , Jan. 2020 DOI: 10.2139/ssrn.3573877
- [11] Banerjee, S. ‘Travelling and COVID-19: A Mathematical Model’, April 2020 DOI: 10.21203/rs.3.rs-26244/v1 LicenseCC BY 4.0

- [12] Koczkodaj, W.W et , 1000,000 cases of COVID-19 outside of China: The date predicted by a simple heuristic, *Global Epidemiology* (2020), <https://doi.org/10.1016/j.gloepi.2020.100023>
- [13] Mazurek, J. and Neničková , Z, Predicting the number of total COVID-19 cases in the USA by a Gompertz curve, April 2020, DOI: 10.13140/RG.2.2.19841.81761
- [14] Mazurek, J. et al, Forecasting the number of total COVID-19 cases and deaths in the World, UK, Russia and Turkey by the Gompertz curve, May2020, DOI: 10.13140/RG.2.2.11336.88321
- [15] Manav R. Bhatnagar, Anubhav Bhatnagar and Supriti Bhatnagar “A Study of COVID19 with Limited Testing Kits,” ResearchGate, DOI: 10.13140/RG.2.2.29138.09923, published on April, 2020.
- [16] Manav R. Bhatnagar, “A Statistical Model for the Spread of COVID19 in Clusters,” ResearchGate, DOI: 10.13140/RG.2.2.18583.52644, published on April, 2020
- [17] Bhatia, V. and Mitra, R. ‘Signal Processing based Predictor for COVID-19 Cases’, April 2020 DOI: 10.13140/RG.2.2.23431.55201
- [18] Arti, M.K., A New Model and Relay Based Study for the Spread of COVID19, <https://www.researchgate.net/publication/340947675>
- [19] Antoni Wilinski, "COVID-19: Model for the spread of the epidemic in a given country allowing determining the phase of its advancement", May 2021 DOI: 10.13140/RG.2.2.26951.42403 (Research Gate)
- [20] A. Jaglan, D. Trehan, Megha, and P. Singhal, "COVID-19 Trend Analysis using Machine Learning Techniques", *International Journal of Scientific and Engineering Research*, 11(12):1162-1167, Dec. 2020.
- [21] Douglas A. Reynolds. Gaussian Mixture Models, Jan. 2008, DOI:10.1007/978-0-387-73003-5-196 (Research Gate)
- [22] Bracher, J., Wolfram, D., Deuschel, J., Goergen, K., Ketterer, J. L., Ullrich, A., ... & Schienle, M. (2020). Short-term forecasting of COVID-19 in Germany and Poland during the second wave—a preregistered study. *medRxiv*.
- [23] Chowdhury, R., Heng, K., Shawon, M. S. R., Goh, G., Okonofua, D., Ochoa-Rosales, C., ... & Franco, O. H. (2020). Dynamic interventions to control COVID-19 pandemic: a multivariate prediction modelling study comparing 16 worldwide countries. *European journal of epidemiology*, 35(5), 389-399.
- [24] Malesza, M., & Kaczmarek, M. C. (2021). Predictors of anxiety during the COVID-19 pandemic in Poland. *Personality and individual differences*, 170, 110419.

- [25] Cherniha, R., & Davydovych, V. (2020). A mathematical model for the COVID-19 outbreak and its applications. *Symmetry*, 12(6), 990.
- [26] Adamczyk, K., Herman, M., Fraczek, J., Piec, R., Szykula-Piec, B., Zaczynski, A., ... & Franek, E. (2020). Sensitivity and specificity of prediction models based on gustatory disorders in diagnosing COVID-19 patients: a case-control study. *medRxiv*.
- [27] Maziarz, M., & Zach, M. (2020). Agent-based modelling for SARS-CoV-2 epidemic prediction and intervention assessment: A methodological appraisal. *Journal of Evaluation in Clinical Practice*, 26(5), 1352-1360.
- [28] Wieczorek, M., Siłka, J., & Woźniak, M. (2020). Neural network powered COVID-19 spread forecasting model. *Chaos, Solitons & Fractals*, 140, 110203.
- [29] Lalmuanawma, S., Hussain, J., & Chhakchhuak, L. (2020). Applications of machine learning and artificial intelligence for Covid-19 (SARS-CoV-2) pandemic: A review. *Chaos, Solitons & Fractals*, 139, 110059.
- [30] Bhardwaj, R., & Agrawal, A. (2021). Analysis of Second Wave of COVID-19 in Different Countries. *Transactions of the Indian National Academy of Engineering*, 1-7.
- [31] Scheiner, S., Ukaj, N., & Hellmich, C. (2020). Mathematical modeling of COVID-19 fatality trends: Death kinetics law versus infection-to-death delay rule. *Chaos, Solitons & Fractals*, 136, 109891.
- [32] Frank, T. D. (2020). COVID-19 interventions in some European countries induced bifurcations stabilizing low death states against high death states: An eigenvalue analysis based on the order parameter concept of synergetics. *Chaos, Solitons & Fractals*, 140, 110194.
- [33] Wilinski, A., & Szwarc, E. (2021). A classification of countries and regions by degree of the spread of coronavirus based on statistical criteria. *Expert systems with applications*, 172, 114654.
- [34] www.pokazwirusa.pl/wykresy [Accessed on July 14]