

# Stochastic Modelling for Coronavirus (COVID'19) Pandemic in Bulgaria

Maroussia Slavtchova-Bojkova<sup>1,2</sup> and Valeriya Simeonova<sup>1</sup>

<sup>1</sup> Faculty of Mathematics and Informatics, Sofia University, No5, J. Bourchier Blvd., 1164 Sofia, Bulgaria

<sup>2</sup> Institute of Mathematics and Informatics. Bulgarian Academy of Sciences

## Abstract

The aim of this paper is to present the development and improvements done in the specific stochastic branching model during the progress of the COVID'19 pandemic caused by SARS-CoV-2 coronavirus up to spring of the year 2022. Our approach is data-driven and uses the parsimonious continuous time Crump-Mode-Jagers branching processes (CMJBP) model. The model provides a basis for decision makers to understand the likely trade-offs as an outbreak begins.

## Keywords

SARS-CoV-2 coronavirus, immigration, Crump-Mode-Jagers branching processes

## 1. Introduction

The theory of branching processes is a powerful tool for study of population dynamics where the members of the population produce new members whose evolution is driven by one and the same stochastic laws. That is why branching processes have relevant applications in physics, biology, medicine, demography, epidemiology, economics, computer science and many other areas where such patterns appear in a natural way. We would like to point out the classical monograph [1] and some recent ones [2]–[6] among others where fundamental models and results of branching processes are presented. Related to the applications of branching processes in biology and medicine [7]–[9] could be pointed out as useful references. On the other hand, statistical inferences of branching processes are considered in [10]–[13].

This paper continues the stochastic modelling research based on the continuous time Crump-Mode-Jagers branching model started in [14] at the beginning of the COVID'19 pandemic. In the two-year period of time we were witnessed of

Information Systems & Grid Technologies: Fifteenth International Conference ISGT'2022, May 27–28, 2022, Sofia, Bulgaria  
EMAIL: bojkova@fmi.uni-sofia.bg (M. Slavtchova-Bojkova); simeonova@fmi.uni-sofia.bg (V. Simeonova)  
ORCID: 0000-0002-8832-7336 (M. Slavtchova-Bojkova); 0000-0002-6515-169X (V. Simeonova)



five waves of the pandemic due to the different variants of the SARS-CoV-2 coronavirus, which were spreading with various intensities all over the world. That is leading to the need of many improvements and modifications of the general tool [15] and in [16] we are using for the simulation and computational procedures, as well.

In what follows we would like to make a short review of the papers devoted to the stochastic modelling of COVID'19 in Bulgaria for the same period. Concerning the lines of research towards the modeling of COVID'19 and statistical procedures for estimating model parameters, the recent paper by [17] could be pointed out. The results reported there extend the previous research of the authors initiated in [18]. More precisely, the authors incorporated into the model an immigration component, vaccination policy and adaptive immunity as additional factors that turned out to be important for the pandemic spreading in Bulgaria. The parameters of the model were estimated from daily available data, representing the number of registered infected individuals. That is mainly the basic reproduction number or the mean number of daily secondary cases, i.e., number of infected individuals by one individual per day. The analyses are made in the frame of two-type Galton-Watson branching process.

In [19] using official data for Bulgaria, the authors investigated the relation between registered COVID'19 infections and related deaths by age and time to estimate the absolute and relative mortality risk. Their aim is to design a basic framework for risk management strategies in order to minimize deaths during the ongoing COVID'19 epidemic in countries with similar quality of healthcare. The used statistical methods are based on dynamic regression models and classical time series analysis.

The paper is organized as follows: Section 2 briefly reminds some preliminaries of the CMJBP model, while Section 3 is devoted to the improvements of the statistical methodology developed in [16] and aims to present the results established by simulation method with especially calibrated parameters of the model fitted to the available data.

We end up the paper by discussion of the results and a brief plan for future work in Section 4.

## **2. Preliminaries of Crump-Mode-Jagers branching processes model**

Before proceeding we give outline descriptions of some common branching process models (see e.g., Jagers [9] for further details), which describe the evolution of a single-type population, which in what follows will be supposed to be the one of infected individuals. In all these branching models, individuals have independent and identically distributed reproduction processes. The re-

production process in terms of epidemic spread is meaning the random process signifying the new infected by each contact with infectious one. In the case of SARS-CoV-2 coronavirus it is known that each contact results in new infective. In a general branching process, also called a Crump-Mode-Jagers branching process (CMJBP), each individual lives until a random age, distributed according to  $I$ , and reproduces at ages according to a point process  $\zeta$ . More precisely, if an individual,  $i$  say having reproduction profile  $(I, \zeta_i)$ , is born at time  $b_i$  and  $0 \leq \tau_{i1} \leq \tau_{i2} \leq \dots \leq I_i$  denote the points of  $\zeta_i$ , then individual  $i$  has one child at each of times  $b_i + \tau_{i1}, b_i + \tau_{i2}, \dots$ . This model allows that a mother could have more than one child during her life or in terms of epidemic that every contaminated case could contact and pass the viral infection to more than one susceptible during its infectious period. However, the situation with SARS-CoV-2 coronavirus is rather different in comparison to other viruses existed till now. It was reported that an individual could just transfer the virus without being ill and/or symptomatic, which complicates the contact process as a whole and the tracing of contacts consequently.

This paper is primarily concerned with models for epidemics of diseases which follow the so-called SEIR (Susceptible  $\rightarrow$  Exposed  $\rightarrow$  Infective  $\rightarrow$  Removed) scheme in a closed, homogeneously mixing population or some of its extensions. A key epidemiological parameter for such an epidemic model is the basic reproduction number  $R_0$  (see Heesterbeek and Dietz [20]), which in the present setting is given by the mean of the offspring distribution of the approximating branching process. A major outbreak (i.e., one whose size is of the same order as the population size) occurs with non-zero probability if and only if  $R_0 > 1$ .

Let us point out another interesting model proposed in [21] where a gamma negative binomial branching process is accepted for the number of new infections generated by an infected individual.

Suppose that  $R_0 > 1$  and some preventive transmission measures are taken in advance of an epidemic. If there were a vaccine this could be expressed in such a way that fraction  $c$  of the population is vaccinated with a perfect vaccine in advance of an epidemic. Then  $R_0$  is reduced to  $(1 - c) R_0$ , since a proportion  $c$  of infectious contacts is with vaccinated individuals. It follows that a major outbreak is almost surely prevented if and only if  $c \geq 1 - R_0^{-1}$ . This well-known result, which gives the critical vaccination coverage to prevent a major outbreak and goes back at least to 1964 (e.g., Smith [22]), is widely used to inform public health authorities, but only if there is a vaccine.

### 3. Improvements of the statistical method and simulation results

#### 3.1. Improvements of the statistical modelling

Our methodology is based primarily on the CMJBP as a model of epidemic spread. We would like to point out here the fundamental result of Ball and Donnelly [23] giving us the theoretical evidence of the almost sure convergence of the epidemic models to the proper CMJBP under certain conditions. By use of the statistical software especially developed for branching processes simulations [16] we first fit the parameters of the model to the data available for particular country. The two main characteristics running the behavior of the CMJBP are the distribution of the life period of individuals and the point process governing the reproduction process of any infected individual which may depend on the age of individual. These quantities in terms of epidemic spreading mean the distribution of the serial interval which is the sum of incubation period and delay period and the point process signifying the number of new infected individuals any infective individual may pass the virus to.

In the present study for the parameters of the CMJBP, we use the left-truncated normal distribution  $N(35, 5.12)$ , using the estimates from [16]. For the point process modelling, i.e. the number of infected individuals by one infected, we use gamma distribution with appropriately defined parameters, i.e.  $\Gamma(7.27, 1.32)$  (see [16]).

The essential changes we make in the tool [15] are as follows: we started with variables' initialization: dates range,  $R_0$ , horizon,  $\mu$  – the immigration mean, confidence interval, number of simulations. We have applied three forecast models called:

- Main scenario corresponding to the case when an immigration and  $R_0$  estimation is determined, so that the algorithm recalculates them until is found such  $R_0$  s in  $[0.3; 5]$  which are smoothly changing in time. For them is calculated the immigration mean (outside incoming infectious people), being the newly infected ones in the population;
- Optimistic scenario that is the main scenario with new  $\tilde{R} = R_0 - 0.2$ ;
- Pessimistic scenario that is the main scenario with new  $R_0 + [0.8; 1.2]$ .

Data filtering options have been added so that process research can be done after a certain date. They not to be dependable by year (it was hardcoded in the original version). We have added the ability to smooth the data, given that we are seeing several waves of the pandemic. We have created a version of the code with Bulgarian text on the generated graphics (before is missing language support). The number of simulations initially in [15] was set to be 1000. It was time consuming still in parallel work. After repeated attempts and testing with the number of simulations in the interval  $[100; 1000]$  it was found that the critical minimum

that provides normally good and almost indistinguishable results compared to 1000 is 150 simulations. Therefore, all subsequent predictions were made with 150 simulations.

### 3.2. From the last two waves in Bulgaria

We are illustrating the methodology using the CMJBP after fitting the theoretical model to the historical data published at Worldometer (see [25]) for Bulgaria illustrated the fourth wave on Figure 1 and fifth one on Figure 2. This way we acquire the values which are best revealing and explaining the structure of the historical data representing the new cases and total cases, as well. Then we are projecting further the behavior of the new daily cases in three scenarios in 45 days horizon.

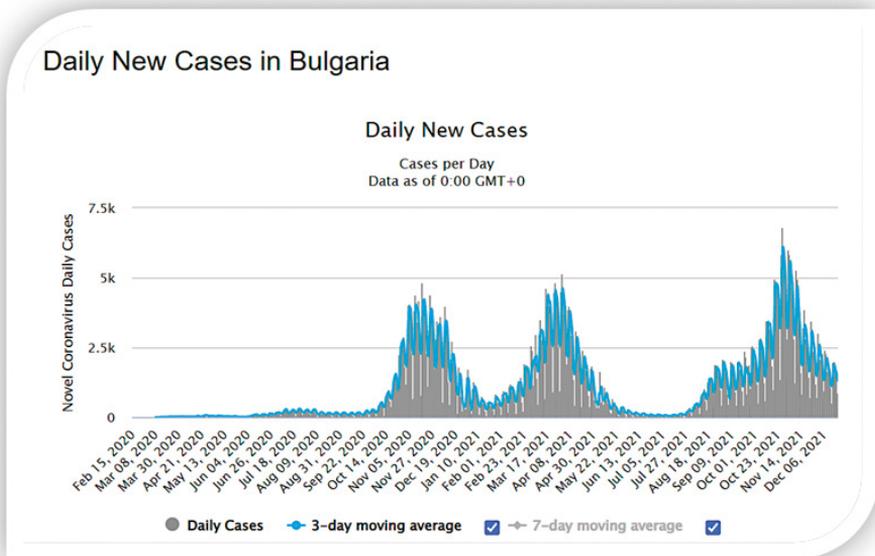
#### 3.2.1. Fourth wave caused by the Delta variant of the virus

We have considered the period between July 27, 2021 – December 12, 2021, as the fourth wave of pandemic with observed maximum on October 26, 2021, of new cases of 6 816 and duration of 5 months.

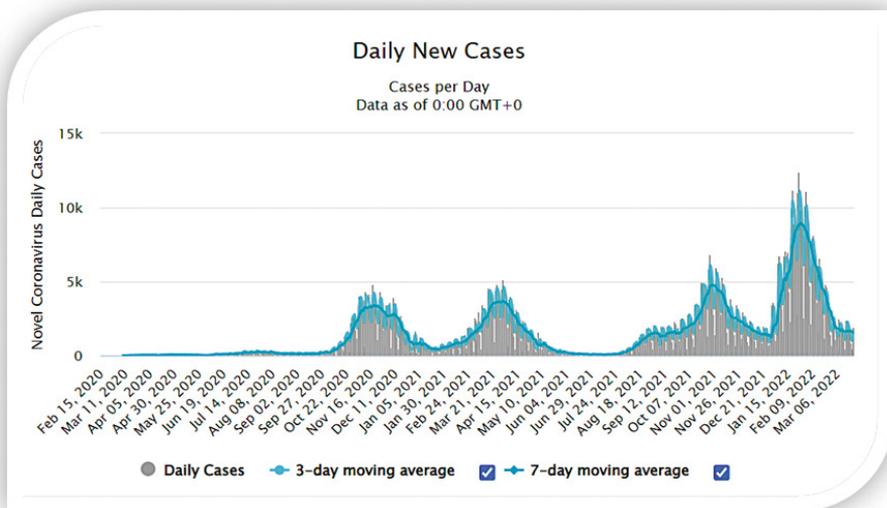
We have obtained the following computer visualization as a result of the implementation of the code:

- the model fit versus observed data for the total and new daily cases during the period between July 27, 2021 –December 12, 2021 (see Figure 3);
- the estimated  $R_0$  and immigration (see Figure 4);
- the three forecasts for new daily cases corresponding to the three scenarios (see Figure 5).
- On Figure 3 (on the left) could compare the behavior of the branching processes model (in blue dotted line) to the observed size (in black line) for the total cases and new daily cases during the fourth wave. The conclusion that could be drawn is that the graphs reveal the good fit of the model to data observed, in general.

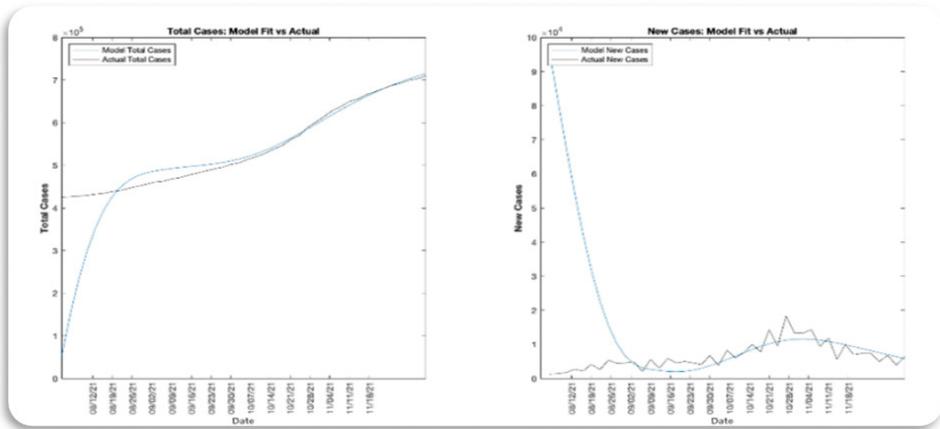
It is worth mentioning here (Figure 4) that the peak of curve representing the behavior of  $R_0$  has its peak in the week between 21st and 28th of October 2021, which is agree with the peak in real data observed on 26th of October. On the other side, the immigration curve (in red dotted line) is following the occurrence of immigration wave at the beginning of August 2021, then its gradual decrement and slight increment after that at the end of the year.



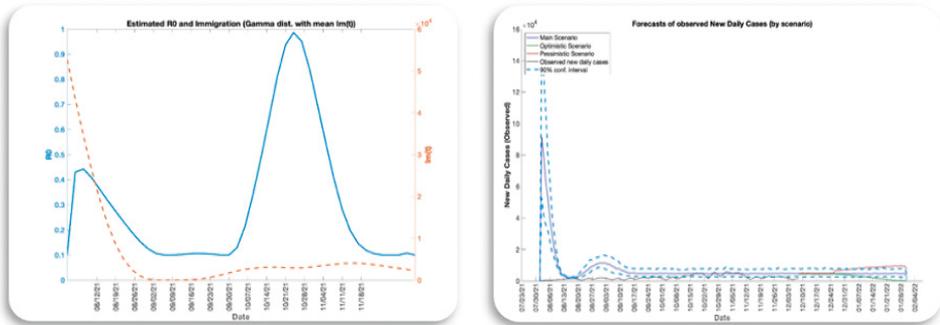
**Figure 1:** The fourth wave of the pandemic in Bulgaria – July 27, 2021 – December 12, 2021



**Figure 2:** The fifth wave of the pandemic in Bulgaria – 15 December 2021 – 25 March 2022



**Figure 3:** Fit of the model vs observed data



**Figure 4:** The change of estimated basic reproduction number and immigration in time

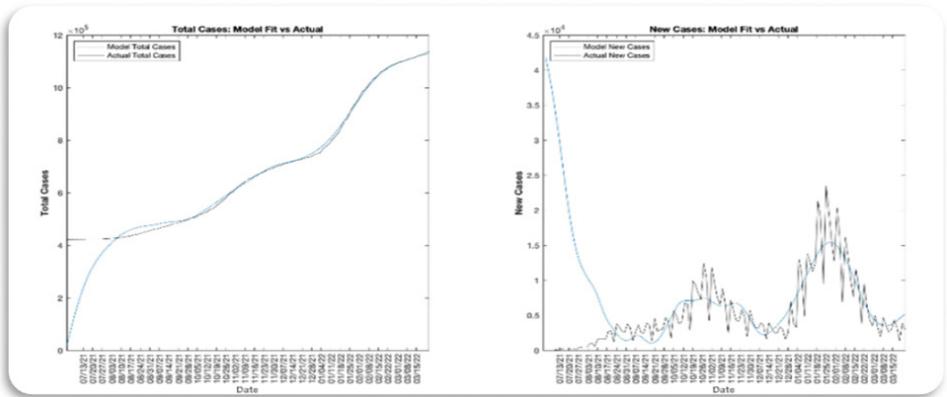
**Figure 5:** Forecast of new daily in Bulgaria – 45 days forecasts

### 3.2.2. Fifth wave caused by the Omicron variant of the virus

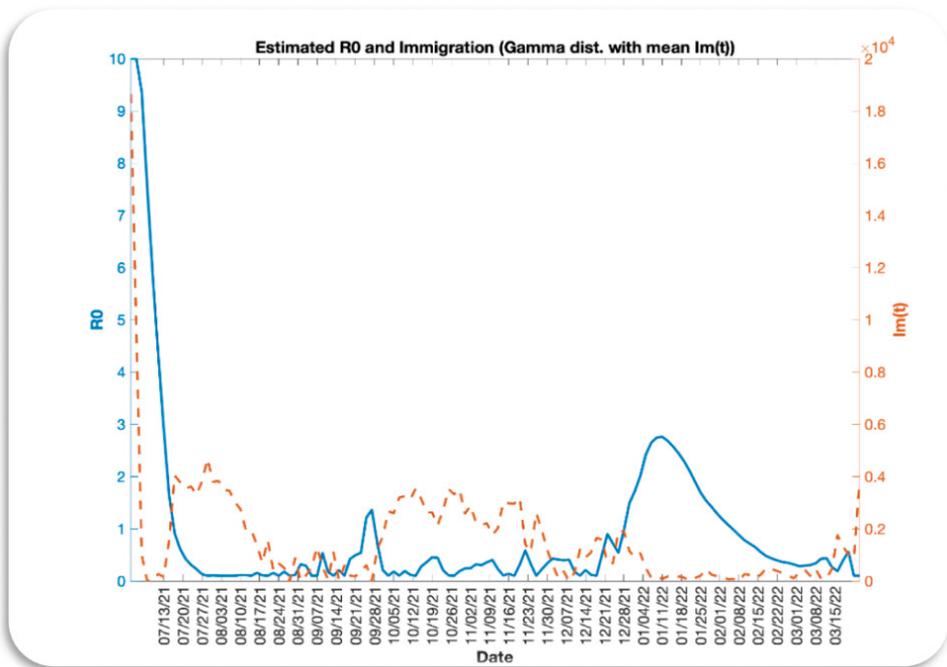
We have considered the period between December 15, 2021 – March 25, 2022 as the fifth wave of pandemic with observed maximum on January 25, 2022 of new cases of 12 399 and duration of 2.5 months.

We have obtained the following computer visualization as a result of the implementation of the code:

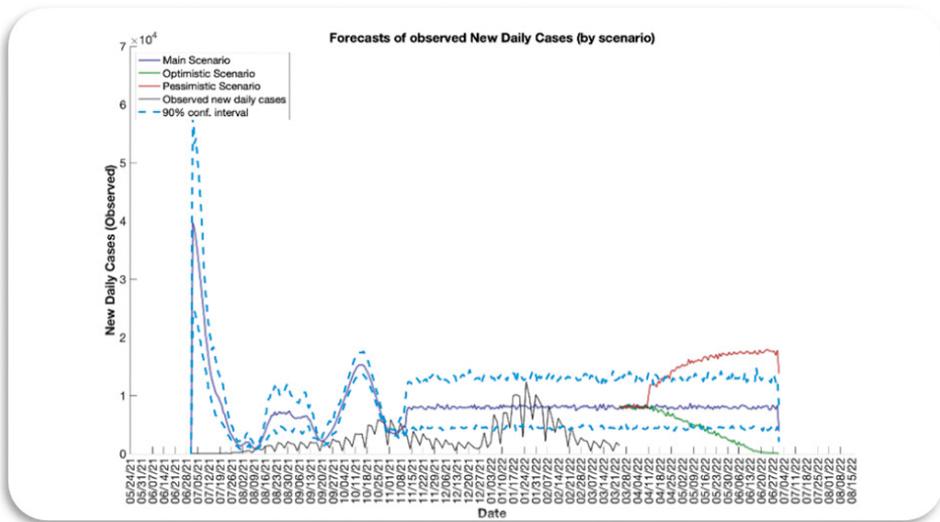
- the model fit versus observed data for the total and new daily cases during the period between July 13, 2021 – March 25, 2022 (see Figure 6);
- the estimated  $R_0$  and immigration (see Figure 7);
- the three forecasts for new daily cases corresponding to the three scenarios (see Figure 8).



**Figure 6:** Fit of the model vs observed data during the period between



**Figure 7:** The change of estimated basic reproduction number and immigration in time



**Figure 8:** Forecast of new daily in Bulgaria – 45 days forecasts

The Fifth wave was not possible to be calculated accurately before parameters’ changes in MATLAB tool, described in Sections 3.1 and 4.1, nor just the Fifth wave, nor both Fourth and Fifth waves. The reason in Figure 6, Figure 7, and Figure 8 to be shown both waves is to test if the tool has the possibility to analyze and fit both waves. It is evident the success we made with the parameters’ optimization. The optimization shows correctly at the same time main, optimistic, and pessimistic scenario including in retrospection. If we have to make some general remark on the obtained results, we could say there is no significant difference in trends for all scenarios in both waves. We could make similar conclusions also for the Immigration processes, shown on Figure 4 and Figure 7.

## 4. Discussion

### 4.1. Discussion on obtained prognosis with branching processes

In this paper we have presented an updated mathematical tool to tackle pandemic outbreaks considering the effect of immigration. This tool (see description in 3.1) addresses various technical questions to support the ongoing public health response to COVID’19. Our approach considers both estimation efforts for key parameters, and investigative efforts (often numerical simulations) in assessing the effectiveness of various intervention or control measures. The tool is written in MATLAB and has the described structure below:

- `buildPlots_filtered_plots_wit_bg_subs.m`: includes only function `buildPlots(NewCasesDaily, TotalCasesDaily, ActiveCasesDaily, NewCase-sHist, dates, horizon, detection_time, alpha, scenario_name, scenario_title, varargin)` which is used for plot generating with Bulgarian interface
- `buildPlots.m`: the original plot generating function
- `Confinterval.m`: includes only one function that calculates the confidence interval: function `[Z_mean, Z_lower, Z_upper, Z_median]=confInterval(Z, alpha)`. Currently, it works with the mean measure, but it is possible to be used the calculated median with branching process's function.
- `Im_function.m`: The function determines immigration process: `Im_matrix = Im_function(mu, sgm, Im_age_struct_prob)`
- `obj_Mu_and_Im.m`: The function aims to evaluate BP's parameters such as  $R_0$  and the Immigration probability.
- `Y_projection.m`: Applies numerical method to calculate the BP's generation with the presumption there is no fatal cases at any time point. This should be used by BP Simulator.
- `readCSVData.m`: used for CSV data reading. We rewrote the script as the time pass the resource links were changed several times.
- `readWorldometerData.m`: Reads COVID'19 data from Worldometer. We rewrote the main part of the script, so now it is not dependent by the year.
- `BranchingProcessSimulator.m`: Includes the function that simulates Branching processes.
- `new_bg.m`: The main script. It includes parameters' settings and uses all the other script. We modified the script, so now it is possible to choose date range and time lags.

It is possible the tool to be optimized using Python and fully rewritten including in additionally graphic user interface. Even more BP model is applicable not only in analyzing epidemic data but also in analyzing demographic processes and many other areas, that implying branching in next generations.

## 4.2. Comparative analysis with other forecasting Methods

Some of the classical methods for forecasting time forecasts with variables have been selected: Autoregression (AR), Moving Average (MA), Autoregressive Moving Average (ARMA), Autoregressive Integrated Moving Average (ARIMA), Seasonal Autoregressive Integrated Moving-Average (SARIMA), Simple Exponential Smoothing (SES), Holt Winter's Exponential Smoothing (HWES). These methods are widely used for data analyzing when it is needed to forecast a process.

Due to the nature of time-series data, there are several specificities involved in time series modeling that are not relevant to other datasets as the timestamp

that identifies the data has intrinsic meaning. Time-series data are mostly Univariate, which doesn't mean there is no Decomposition. Contrariwise, the decomposition is a technique to determine whether there is seasonality, trend or stationery, and noise in data. The main rule is if there is present trend then the data are not stationary. The time series models bellow are not able to deal with trends. We can detect non-stationarity using the Dickey-Fuller Test and if it is present, we can remove non-stationarity using differencing.

ARIMA family (AR, MA, ARMA, ARIMA, SARIMA, SARIMAX) models are used to forecast the observation at (t+1) based on the historical data of previous time spots recorded for the same observation. All these models give close enough prediction about any particular time series.

Simple research in Web of Science shows what looks like the use or ARIMA family in forecasting over the years, as all numbers are in terms of 'about':

Applied model	ARMA	ARIMA	SARIMA
Over the time	5000	9000	1100
In last 6 years	1500	4000	600
In last 6 years, only for analyzing COVID'19 data (only by search term 'covid')	23	345	44

We can make an aggregated formula for all ARIMA family models in form of parameters describing the model:

*ARIMA\_family\_model (p,d,q)*

will describe models AR, MA, ARMA, ARIMA, SARIMA, where:

**p:** is how many previous timesteps adjusted by a multiplier and then adding white noise we are using. Comes from AR model.

**d:** is the difference order, which is the number of transformations needed to make the data stationary. Comes from I model.

**q:** is the number of lagged forecasting error terms in the prediction. Comes from MA model. In an MA (1) model, the forecast is a constant term plus the previous white noise term times a multiplier, added with the current white noise term. This is just simple probability and statistics, as we are adjusting our forecast based on previous white noise terms.

Obtained result for stationarity has p-value<0.05 which means that COVID'19 time-series is stationary, so all the above models are applicable. We are using AR, MA, ARMA and ARIMA with the presumption the data are not seasonal, as there is not exactly and identical measure for seasons in COVID'19 data nevertheless we have several waves. In the table below are shown the values of parameters p, d, and q for MA, ARMA, ARIMA and SARIMA that were applied testing the model with COVID'19 data:

Model	p	d	q
AR	1	NA	NA
MA	NA	NA	1
ARMA	2	0	1
ARIMA	1	1	1
SARIMA	1	1	1

SARIMA model is very similar to the ARIMA, except that there is an additional set of autoregressive and moving average components. The additional lags are offset by the frequency of seasonality (ex. 12 — monthly, 24 — hourly). Implementation doesn't include seasons, but it is possible to introduce for example weekly dependence as it is a well-known fact the values for Saturday and Sunday often are smaller than the values from the other days. So, on a weekly base it is possible to identify mini waves. The other opportunity is to approximate the intervals between picks of waves and to use this measure as seasonal parameter.

Data on new daily cases for COVID'19 were downloaded from the Worldometer website and processed using the listed methods using Python [27] with Python module *statsmodels 0.14.0* [28]. There is comparability because the algorithm for branching processes uses the same data set. Forecasts are generated quickly and represent specific integers that should reflect new cases. Each predicted value is added to the original data so that we can predict the next one in chronological order. The forecast horizon is 45 days.

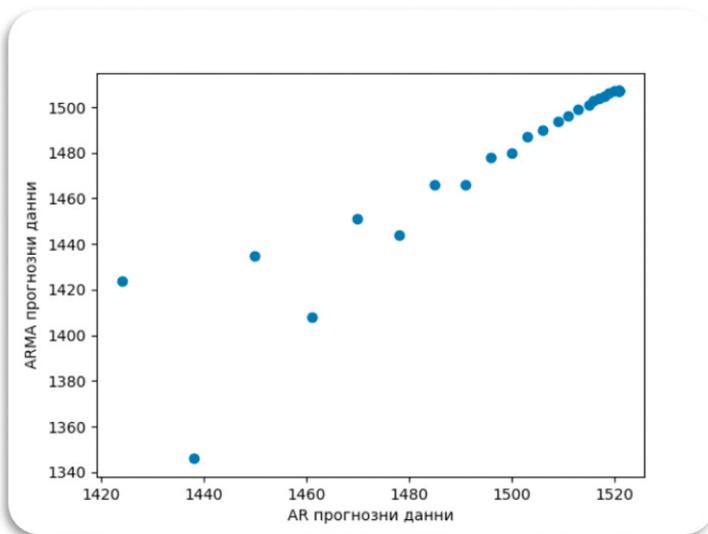
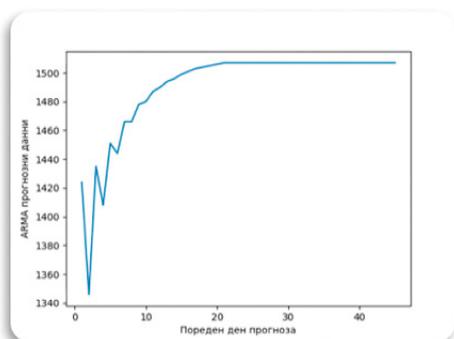
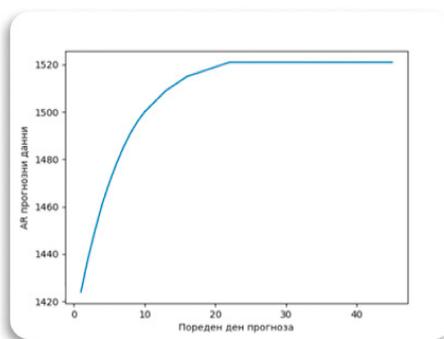


Figure 9: ARMA vs ARIMA prognosis



**Figure 10: ARIMA forecast**



**Figure 11: AR forecast**

The SES and HWES methods are not applicable, as manage to make a forecast in just two days and then there is no change in the forecast value within the desired forecast period. The ARIMA and SARIMA methods give similar results, with the difference that they manage to make a forecast for the next 6 days. However, they could be used for short-term forecasts, and in both cases the trend is sharply rising, but the slope is low. The results themselves are identical. The AR and ARMA methods are interesting because the forecast is not monotonous. For them, the maximum of possible days for forecast is 26, and then continues unchanged. We could refer them to the pessimistic variant in the simulation of branching processes. The MA method gives a negative trend but manages to make predictions for 3 consecutive days. It could be assumed that it reflects the optimistic option in simulating branching processes.

The conclusion to be drawn is that branching processes can be taken as a method that can summarize the trends, we obtain through other forecasting methods.

### 4.3. Future work

Work on the analysis of the predictions obtained through branching processes will continue with the conduct of a series of comparative analyzes, including modern methods such as neural networks, including transfer neural networks. At a minimum, the following forecasting options will be considered:

1. A neural network that uses the entire time series of data
2. A set of at least 3 separate neural networks, each working on a piece of data. The idea comes from the fact that the algorithm we use to predict branching processes makes data smoothing – sometimes every 3 days, sometimes every 5 days. This allows us to use different windows to generate different data sets based on the original complete data set. This technique will ensure, among other things, comparability of results with different start dates and the same smoothing window.

3. Variant of 2., in which neural networks will be represented as a convolutional neural network. The result should be a consolidated forecast with a horizon of 45 days.

4. Variant of 2. using transfer training. A study by Todor Tsokov, Milena Lazarova and Adelina Aleksieva-Petrova will be used for the “Spatial-temporal neural model for predicting air pollution”, presented at the spring scientific session of FMI 2022 FMI SU “St. Kliment Ohridski” [26]. We believe that there is a great similarity between the spread of dust particles and the spread of viruses. This gives us reason to test the theory by applying transfer learning with the data from the distribution of COVID’19 on the neural network from the cited study.

## 5. Acknowledgements

This work was supported by the project BG05M2OP001-1.001-0004 (UNITE) funded by Operational Program Science and Education for Smart Growth co-funded by European Regional Development Fund for theoretical development of the stochastic modelling of COVID’19 pandemic by means of branching processes and for computational resources by the project KP-6-H22/3 of NSF at the Bulgarian Ministry of Education and Science.

## 6. References

- [1] K. B. Athreya and P. Ney, Branching processes. Springer-Verlag Berlin, New York, 1972.
- [2] M. González Velasco, I. M. del Puerto García, and G. P. Yanev, Controlled branching processes.
- [3] P. Haccou, P. Jagers, and V. A. Vatutin, Branching Processes. Cambridge University Press, 2005.
- [4] T. E. Harris, The Theory of Branching Processes. Berlin, Heidelberg: Springer Berlin Heidelberg, 1963.
- [5] C. J. Mode, Multitype branching processes : theory and applications. New York (N.Y.) : American Elsevier, 1971.
- [6] B. A. Sevast’yanov, Ветвящиеся процессы. (Russian) [Branching processes]. Moscow: Nauka, 1971.
- [7] A. Y. Yakovlev and N. M. Yanev, Transient Processes in Cell Proliferation Kinetics, vol. 82. Berlin, Heidelberg: Springer Berlin Heidelberg, 1989.
- [8] M. Kimmel and D. E. Axelrod, “Branching processes in biology,” in *Interdisciplinary Applied Mathematics*, 2<sup>nd</sup> ed., vol. 19, Springer, 2015, pp. 1–280.
- [9] P. Jagers, ‘Branching Processes with Biological Applications’. John Wiley & Sons, London-New York-Sydney-Toronto 1975. VIII, 268

- S.,” *Biometrical J.*, vol. 20, no. 1, pp. 94–94, Jan. 1978, doi: 10.1002/bimj.4710200112.
- [10] P. Guttorp, *Statistical inference for branching processes*. New York: Wiley, 1991.
- [11] V. Stoimenova, D. Atanasov, and N. Yanev, “Robust Estimation and Simulation of Branching Processes,” *Comptes Rendus l’Academie Bulg. des Sci.*, vol. 57, p. 5, Jan. 2004.
- [12] A. Y. Yakovlev, V. K. Stoimenova, and N. M. Yanev, “Branching Processes as Models of Progenitor Cell Populations and Estimation of the Offspring Distributions,” *J. Am. Stat. Assoc.*, vol. 103, no. 484, pp. 1357–1366, Dec. 2008, doi: 10.1198/016214508000000913.
- [13] N. M. Yanev, “Statistical inference for branching processes,” *Rec. Branch. Process.*, pp. 143–168, 2008.
- [14] P. Trayanov, “MATLAB Branching Process Simulator.” GitHub, 2020, [Online]. Available: <https://github.com/plamentrayanov/BranchingProcessSimulator>.
- [15] P. Trayanov, “MATLAB fitCOVID19\_BranchingProcess.” GitHub, 2020, [Online]. Available: [https://github.com/plamentrayanov/fitCOVID19\\_BranchingProcess/releases/tag/v1.0.1](https://github.com/plamentrayanov/fitCOVID19_BranchingProcess/releases/tag/v1.0.1).
- [16] M. Slavtchova-Bojkova, “Branching processes modelling for coronavirus (COVID’19) pandemic,” in *CEUR Workshop Proceedings*, 2020, vol. 2656, pp. 115–130, Accessed: Apr. 18, 2022. [Online]. Available: <http://ceur-ws.org/Vol-2656/paper12.pdf>.
- [17] D. Atanasov, V. Stoimenova, and N. M. Yanev, “Branching Process Modelling of COVID-19 Pandemic Including Immunity and Vaccination,” *Stochastics Qual. Control*, vol. 36, no. 2, pp. 157–164, Dec. 2021, doi: 10.1515/eqc-2021-0040.
- [18] N. M. Yanev, V. K. Stoimenova, and D. V. Atanasov, “Stochastic modelling and estimation of COVID-19 population dynamics,” *Comptes Rendus L’Academie Bulg. des Sci.*, vol. 73, no. 4, pp. 451–460, Jun. 2020, doi: 10.7546/CRABS.2020.04.02.
- [19] L. Tomov, A. Tchorbadjieff, and S. Angelov, “Age-specific mortality risk from Covid-19 in Bulgaria.” 2021, doi: 10.13140/RG.2.2.26678.01600/1.
- [20] J. A. P. Heesterbeek and K. Dietz, “The concept of  $R_0$  in epidemic theory,” *Stat. Neerl.*, vol. 50, no. 1, pp. 89–110, Mar. 1996, doi: 10.1111/j.1467-9574.1996.tb01482.x.
- [21] [J. Levesque, D. W. Maybury, and R. H. A. D. Shaw, “A model of COVID-19 propagation based on a gamma subordinated negative binomial branching process,” *medRxiv*, p. 2020.07.08.20149039, Oct. 2020, doi: 10.1101/2020.07.08.20149039.

- [22] C. E. Smith, “Factors in the transmission of virus infections from animals to man.,” *Sci. Basis Med. Annu. Rev.*, pp. 125–150, 1964.
- [23] F. Ball and P. Donnelly, “Strong approximations for epidemic models,” *Stoch. Process. their Appl.*, vol. 55, no. 1, pp. 1–21, 1995, [Online]. Available: <https://econpapers.repec.org/RePEc:eee:spapps:v:55:y:1995:i:1:p:1-21>.
- [24] J. Hellewell et al., “Feasibility of controlling 2019-nCoV outbreaks by isolation of cases and contacts,” *medRxiv*, p. 2020.02.08.20021162, Feb. 2020, doi: 10.1101/2020.02.08.20021162.
- [25] “COVID Live – Coronavirus Statistics – Worldometer.” <https://www.worldometers.info/coronavirus/#countries> (accessed Apr. 17, 2022).
- [26] T. Tsokov, M. Lazarova, and A. Aleksieva-Petrova, “Spatial-temporal neural model for predicting air pollution,” Sofia, 2022. [Online]. Available: <https://intranet.fmi.uni-sofia.bg/index.php/s/W3ynwaJ4zM6dVel>.
- [27] Van Rossum, G. & Drake, F.L., 2009. *Python 3 Reference Manual*, Scotts Valley, CA: CreateSpace.
- [28] Seabold, S. & Perktold, J., 2010. *statsmodels: Econometric and statistical modeling with python*. In 9<sup>th</sup> Python in Science Conference.