A
Review Paper
on
SEIR Modeling of the Italian Epidemic ofSARS-CoV2Using ComputationalSwarm Intelligence

Original Article Authors: Alberto Godio *, Francesca Pace and Andrea Vergnano, Department of Environment, Land and Infrastructure Engineering (DIATI), Politecnico di Torino, Corso Ducadegli Abruzzi 24, 10129 Torino, Italy; francesca.pace@polito.it (F.P.); andrea.vergnano@polito.it (A.V.)*Correspondence: alberto.godio@polito.it; Tel.: +39-011-090-7656

Published in: International Journal of Environmental Research and Public Health

Maheshwari Vidushi
Student, Global Public School
Rajasthan. IND-324009
E-Mail: vidushi.maheshwari12@gmail.com

Kumar Ashutosh
Asst. Professor,
School of Commerce and Management,
Career Point University, Kota-Raj. IN.
E-Mail: ashutoshk.rao@gmail.com

Received: 3 April 2020; Accepted: 15 May 2020;Published: 18 May 2020

Research Article Reviewer/s:

Abstract: In this research article, researchers have applied a generalized SEIR epidemiologicalmodel to the recent SARS-CoV-2 outbreak in the world, with a focus on Italy and its Lombardy, Piedmont, and Veneto regions. This article has been sectioned in 5 titles: introduction, material and methods, results discussion and conclusions. With the help of this article, the research team focused on applying a stochastic approach to fit model parameters using a Particle Swarm Optimization (PSO) solver, to improve the reliability of predictions within 30 days. In the experiment, they also analyzed social data and the expected evolution of the epidemic in Italian regions, and compared the results with data and forecasts from Spain and South Korea. In the next phase, they also linked the model's equations to changes in people's mobility, referring to Google's COVID-19 community mobility reports. Finally, they discussed the effectiveness of the policies adopted by different regions and countries and their impact on past and future infection scenarios.

Keywords:*SARS-CoV-2; COVID-19; SEIR modeling; Italy; stochastic modeling; swarm intelligence*

1. Introduction:

In this researcher article, researchers have presented an updated version of the predictive model of epidemic phenomena based on the approach called SEIR (Susceptible-Exposed-Infective-Recovered), widely used to analyze infection data during the different stages of an epidemic outbreak. The SEIR model represents one of the most widely adopted mathematical models to characterize epidemic dynamics and to predict possible contagion scenarios. The SEIR model can be useful in evaluating the effectiveness of various measures, such as blocking, from an infectious disease outbreak. It is based on a series of dynamic ordinary differential equations that take into account the number of the population prone to infection, the trend over time of individuals recovering after infection, and individuals who sadly die.[1]

This work was carried out during the crucial phase of development of the epidemic in Italy (mid-April 2020), with operational difficulties linked to the impossibility of verifying and validating databases, and with the difficulty of comparing and calibrating the results with other studies. The objective, however, is to provide a useful and easy-to-read tool that can help policy makers, those responsible for strategic decisions, to assess the social and economic scenarios linked to the development of the epidemic.

The generalized SEIR model is based on a system of differential equations, as discussed by Peng et al. (2020)[2] in the analysis of the SARS-CoV-2 epidemic in China. The model, which adds complexity to the classic SIR or SEIR models, represents the different conditions of susceptible and infected individuals during an epidemic (especially quarantined people, who cannot infect others during an epidemic, their quarantine). The coefficients of the equations represent the rates of variation over time of the different categories of individuals, that is, infected, dead and cured.[3]

The main objective of the work is to improve the classic SEIR model by means of a stochastic solver that identifies a set of possible solutions (or most probable scenarios) predicting the evolution of the epidemic with the evaluation of the associated uncertainty. Furthermore, the mathematical model has been modified to adopt a time-dependent infection rate to adequately describe a realistic situation where people's contact is not constant over time due to imposed social distancing rules. The ultimate goal is to provide a reliable approach that predicts the course of the epidemic so that policy makers can undertake both appropriate initiatives to reduce contagion and targeted actions given the uniqueness of each region.

2. Materials and Methods:

2.1. Database

The analysis is based on data collected and available through a control panel at John Hopkins University in the United States. They represent an official database because they collect data from different official bodies such as the World Health Organization (WHO), the European Center for Disease Prevention and Control (ECDC), the Centers for Disease Control and Prevention. from the United States and other organizations. Italian data is collected in its entirety via the Protezione Civile Italiana newsletter. The Italian National Institute of Statistics (ISTAT) is responsible for the number of national and regional populations.The number of people living in the studied regions is reported in Table 1.

---

[1] Parham, P.E.; Michael, E. Outbreak properties of epidemic models: The roles of temporal forcing and stochasticity on pathogen invasion dynamics. J. Theor. Biol. 2011, 271, 1–9.

[2] Peng, L.; Yang,W.; Zhang, D.; Zhuge, C.; Hong, L. Epidemic analysis of COVID-19 in China by dynamical modeling. MedRxiv Epidemiol. 2020.

[3] Bacaër, N. A Short History of Mathematical Population Dynamics; Springer: London, UK, 2011.

Table 1. Population (approximated) for Italy and Italian regions and for other countries included in the following analysis.

Countries/Regions	Overall Population	Database (Year)
Italy	6,03,59,546	Istituto Nazionale di Statistica—ISTAT (2019)
Lombardy	1,00,60,574	Istituto Nazionale di Statistica—ISTAT (2019)
Veneto	49,05,854	Istituto Nazionale di Statistica—ISTAT (2019)
Piedmont	43,56,406	Istituto Nazionale di Statistica—ISTAT (2019)
Spain	4,71,00,396	Istituto Nacional de Estadìstìca—INE (2019)
South Korea	5,16,29,512	Korean Statistical Information Service—KOSIS (Nov. 2018)

2.2 A brief introduction to SEIR Model

Many diseases have a latent phase during which the individual is infected but not yet infectious. This delay between the acquisition of infection and the infectious state can be incorporated within the SIR model by adding a latent/exposed population, E, and letting infected (but not yet infectious) individuals move from S to E and from E to I.

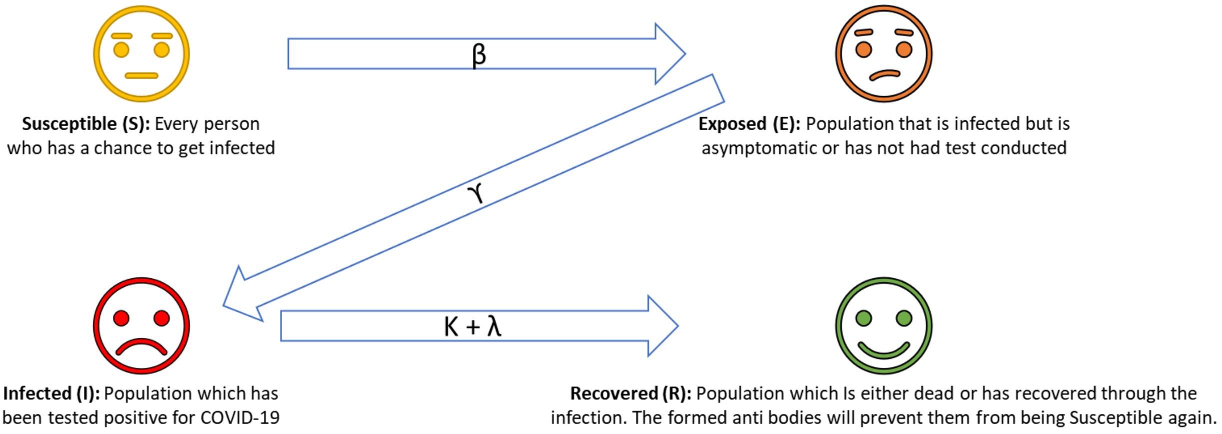

Figure: 1.1 SEIR Model

2.2.1 Overview of the Generalized SEIR Model

The SEIR model simulates the temporal history of an epidemic phenomenon. In its classic form, it models the mutual and dynamic interaction of people between four different conditions, susceptible (S), exposed (E), infectious (I) and cured (R). The classical SEIR model can be described by a series of ordinary differential equations.

$$\frac{dS(t)}{dt} = -\beta \mathrm{I}\ (\mathrm{t})\,.\frac{S(t)}{N}(1)$$

$$\frac{dE(t)}{dt} = \beta \mathrm{I}\ (\mathrm{t})\,.\frac{S(t)}{N} - \Upsilon E(t)(2)$$

$$\frac{dI(t)}{dt} = \Upsilon E(\mathrm{t}) - (\lambda + \kappa)I(t)(3)$$

$$\frac{dR(t)}{dt} = (\lambda + \kappa)I(t)(4)$$

The people who are Susceptible (S) are potential subjects to the virus. Unlike the traditional methods of taking the susceptible population as a whole, this research would be focusing on the part of population which are not yet COVID affected. The Exposed (E) is the fraction of the population that has been infected but does not show symptoms yet, and at this stage, the disease can be infectious, partially infectious or not infectious. The Infective (I) population comprises the transition of exposed population after the time of incubation. The Recovered

(R) population are the people who have moved from the infectious stage. It comprises both dead and recovered populations. The Recovered population can neither advance the disease to others and can neither get infected.

Because of this assumption of closed population, the total population (N) at any time would be:

$$S(t) + E(t) + I(t) + R(t) = N(5)$$

A disadvantage of this method is that researcher not considered the exposed people directly transitioning to the removed population. Here R0 = Reproductive Number

3. Results

After the implementation and use of the previous SEIR model, the researchers presented the time series attained by the standard deterministic approach and the data series obtained by the stochastic approach, based on the (PSO) Particle Swarm Optimization algorithm.

- First, they analyzed the Italian framework at the national and regional level.
- After that they provided the results of the SEIR model for two other countries: Spain and South Korea. Spain was chosen because the spread of the epidemic is look alike to that of Italy, while South Korea represented a test data set because the peak of the epidemic had already passed.
- And the final NMRSE of the modeling and the values of the SEIR coefficients are provided in table 2.

Table 2. Coefficients of the PSO best-solutions and deterministic solutions; bold refers to the best solution of the PSO approach (mean and variance on brackets), ***italic-bold*** refers to the solution given by the deterministic approach.

Country	A	β*	Γ	δ	λ_0	λ_1	κ_0	κ_1	NRMSE
Italy	**0.021** (0.086, 0.004) ***0.012***	**0.510** (1.058, 0.200) ***1.170***	**0.265** (0.859, 0.226) ***1.065***	**0.103** (0.095, 0.01) ***0.020***	**0.017** (0.017, 6.3 x 10^{-9}) ***0.017***	**2** (1.696, 0.180) ***1.983***	**0.029** (0.030, 2.7 x 10^{-6}) ***0.033***	**0.038** (0.040, 3.8 x 10^{-6}) ***0.043***	**0.035** ***0.043***
Spain	**0.037** (0.087, 0.002) ***0.026***	**1.777** (1.376, 0.082) ***2***	**0.946** (0.954, 0.198) ***0.154***	**0.238** (0.095, 0.004) ***0.614***	**0.044** (0.044, 6.8 x 10^{-7}) ***0.043***	**0.156** (0.159, 0.0004) ***0.160***	**0.030** (0.030, 1.6 x 10^{-6}) ***0.028***	**0.046** (0.047, 3.7 x 10^{-6}) ***0.044***	**0.046** ***0.052***
South Korea	**0.292** (0.270 0.0004) ***0.1***	**2** (1.915, 0.009) ***0.974***	**2** (1.846, 0.046) ***1.902***	**0.123** (0.136, 7.8 x 10^{-5}) ***0.313***	**0.05 ****	-	**8.3 x10^{-4}** (8.3 x 10^{-4}, 3.5 x 10^{-11}) ***0.007***	**7 x10^{-6}** (2.1 x 10^{-6}, 2.1 x 10^{-11}) ***0.134***	**0.074** ***0.078***
Lomardy	**0** (0.132, 0.012) ***8.9 x10^{-4}***	**0.460** (1.658, 0.188) ***0.81***	**0.295** (1.093, 0.531) ***0.302***	**0.145** (0.198, 0.065) ***0.253***	**0.027** (0.026, 3 x 10^{-8}) ***0.027***	**0.981** (1.576, 0.247) ***1.925***	**0.036** (0.036, 6.9 x 10^{-6}) ***0.045***	**0.031** (0.031, 6.9 x 10^{-6}) ***0.0405***	**0.062** ***0.061***
Vento	**0.133** (0.102, 0.002) ***0.049***	**1.704** (1.175, 0.190) ***0.97***	**0.920** (0.698, 0.144) ***0.246***	**0.032** (0.034, 0.0004) ***0.09***	**0.049** (0.182, 0.093) ***0.099***	**0.009** (0.008, 0.000) ***0.004***	**0.008** (0.008, 6.4 x 10^{-8}) ***0.009***	**0.0215** (0.021, 1.3 x 10^{-6}) ***0.024***	**0.035** ***0.040***
Piedmont	**0.240** (0.163, 0.009) ***0***	**1.990** (1.518, 0.232) ***0.994***	**0.265** (1.191, 0.374) ***0.195***	**0.012** (0.117, 0.052) ***0.344***	**0.386** (0.309, 0.104) ***0.069***	**0.001** (0.005, 2.5 x 10^{-5}) ***0.007***	**0.019** (0.018, 3.7 x 10^{-6}) ***0.019***	**0.034** (0.031, 1.9 x 10^{-5}) ***0.035***	**0.056** ***0.050***

* The β parameter is time-dependent, as explained in the Methods section. The reported value is the initial value. It decreases around 70% after mid-March, that is, after the national blockade of March 11, with slight differences between the Italian regions. This decrease is much less pronounced in the South Korean equation (around 10%); ** For South Korea, the six coefficients of the λ model are not shown because they represent a different mathematical distribution than other countries, as explained in the Methods section. However, based on fit to real data, we can see that λ gradually increases to 0.5 in mid-March; Figures in bold allow to distinguish the best solution of the PSO method; Figures in italics allow distinguishing the solution from the deterministic approach.

4. Discussion:

They have adopted a generalized SEIR model to provide quantitative information on the complex analysis of the SARS-CoV-2 outbreak, while the disease is still ongoing. The parameters were adjusted in the sense of least squares with a deterministic approach, then with a stochastic approach, using a Particle Swarm Optimization (PSO) algorithm, a novelty in the field of epidemiological studies.

Analysis of the results of the stochastic approach provides information on the selected most probable scenarios of the solutions within 5% of the normalized root mean square (NRMSE) of the best solution. For each area studied, the researchers carried out 50 PSO simulation trials and 5 to 15 trials belonged to the most likely set. Note that the predicted model responses led to a roughly equivalent fit of the data (normalized to the mean value at an L2 norm <0.05). Likely scenarios sometimes presented a wide range of possible solutions due to the inherent framework of the stochastic approach. The different scenarios were carried out thanks to a more in-depth study of the domain of the model space where the solutions are not driven and influenced by the initial estimation of the coefficients of the SEIR model. One of the main limitations of the deterministic approach, on the contrary, is that the results are biased by the selection of the starting point of the model parameters.

The data appears to confirm that, while Lombardy and Piedmont have applied similar approaches to social distancing and store closings, the Veneto strategy has applied a much more proactive effort to limit contagion, through extensive testing of symptomatic cases. and asymptomatic from the beginning, jointly tracking positive potentials. The different actions undertaken by the Regions are well represented in the future evolution of the model, with obvious advantages in an early end to the spread of the infection in Veneto. In fact, the peak of quarantines in Veneto precedes those of Lombardy and Piedmont. The downward curve in Veneto has a steeper trend than that of the other two regions. Furthermore, in Veneto, expected deaths are ten times less and recoveries are five times less than in Lombardy.

Piedmont's behavior faces a particular trend, introducing a recovery period (green curve) in relation to death cases (black curve), since the number of recovered in March is always lower than the deceased. This is because the intersection of the trends of cured cases and cases of death is reached later than the other regions analyzed here, that is, on April 7. This is likely the result of the regional screening policy that only assessed (and counted as confirmed) patients with severe or high-risk symptoms. The high death rate in March was also due to unexpected stress on the health system and a shortage of intensive care units. By contrast, Lombardy and Veneto experienced a higher rate of cured patients in the early stages of the epidemic.

A recent analysis[4] has shown how, according to guidelines from central government public health authorities, Lombardy's actions involved a more conservative approach that focused primarily on symptomatic cases. They also assume that the set of policies adopted in Veneto minimized the burden on hospitals and minimized the risk of spread to medical facilities.

---

[4] Pisano, G.P.; Sadun, R.; Zanini, M. Lessons from Italy’s Response to Coronavirus. Available online: https://hbr.org/2020/03/lessons-from-italys-response-to-coronavirus (accessed on 16 May 2020).

Veneto's strategy attempted to prevent local capillary spread, limiting contagious diseases through additional measures at hot spots of infection in the early stages of the epidemic.

The expected trend in these regions has been controlled by many factors beyond the control of policy makers, including the higher population density of Lombardy and the higher number of cases during the explosion of the crisis. However, the different public health policies in the initial stage of epidemic phenomena have also had an impact, and it seems that adapted capillary actions, as in the example of Veneto, it has achieved better results than applying only a regional blockade. The difference in the approaches can be highlighted by observing that many municipalities or provinces have declared "Red zones", where, due to the high transmission of the infection, additional restrictive measures have been adopted introduced, compared to the rest of the regional territory. In the red zones, different policies acted in response to local epidemiological situations. In contrast, in Piedmont and Lombardy, no red zones have been established, but restrictive measures of individual distancing have been regionally regulated.

According to the proven results of different policies, in the next phase of government policies,

The reopening of businesses and activities should be adapted to local situations, focusing on organization and integration of all figures of the health system. In particular, the central government should require regions to make an effort to provide local epidemiological data in real time, to block only limited areas, while the reopening of businesses at the regional level can be facilitated.

The estimated parameters that regulate the equations of the SEIR model are presented in Table 2.

For the parameters obtained with the stochastic approach, the best solution is indicated in bold, while in in parentheses the mean and variance of the solutions within 5% of the minimum NRMSE in parentheses.

In Table 2, the researchers compared the parameters between different regions and between the stochastic and deterministic approach. Both approaches provided models that fit the observed data with good precision, although the stochastic approach generally had a slightly lower NRMSE.

The parameters calculated with PSO are reported with the value, the mean and the variance of the best solution.

According to the observation of the data, α, β, γ and δ had a high variance due to the inherent variability due to the stochastic approach. Sometimes the best solution is not aligned with the mean value. The λ and κ values, on the other hand, cluster strictly around the mean in almost all PSO solutions, hence the low variance. This is explained considering that, since the number of parameters is greater than the available data series (Q, R and D), the problem is underdetermined, so the stochastic approach can find more than one series of parameters that match the data in an acceptable discrepancy. Then, λ and κ do not show great variability between the most probable scenarios because they govern the equations that correlate Q with R and D, that is, the official data series. Therefore, the estimated λ and κ have always been found in the same region of the solution search space. South Korea and Veneto have the highest recovery rates (λ_1) with values around 0.05, followed by Spain (0.044). This confirms reports that praise the Veneto model, as its administration had the ability to perform tests faster than other Italian regions, and family doctors worked in stronger synergy with health structures. There is also a lower mortality rate (κ_0), probably due to a better efficiency of the health system in treating patients, but also to the greater number of tests. In data and policy, the Veneto region is more like South Korea than other parts of Italy.

These aspects had an impact on the outbreak of the epidemic. The Veneto region is more likely to reach the peak of active cases (Q) before the other regions.Even though the PSO results may seem to provide a wide range for the SEIR parameters, therefore researchers have stressed ontwo important aspects:

- As they have already stated, the problem is underdetermined, so it is preferable to have an acceptable range of values than a unique point value, that could result in being uncertain, as could happen considering only a deterministic approach solution.
- The set of possible predicted scenarios, although related to different solutions with different sets of parameters, are quite similar, thus providing an acceptable level of variability for future predictions.

Although it would be very useful to estimate a narrower range of parameters such as infection rate β or latency time ($1 / \gamma$), this exceeds the objectives of the study, and the issue is being explored. by researchers who also focus on the clinical aspects of the disease.

This model has certain limitations that are summarized to highlight possible needs in the later development of the model.

- They currently do not have enough information to say that after recovery an individual becomes completely immune to the disease, but they made this assumption: the model did not allow the transition from the recovered category to the susceptible category.
- The model does not consider that the exposed category may have a partial infectivity, as described in Shi (2020)[5], nor does it distinguish symptomatic from asymptomatic people, as studied in Shaikh (2020)[6].
- The model does not take into account differences in evidence between different structures of the health system and national policies.

Although the Italian and Spanish data fit well, the fit of the South Korean data has some problems.

This shows that different policies between countries can induce different trends in the spread of the epidemic and that models must be adapted to different situations, with the introduction or remove parameters. This would be particularly valid for analyzing the situation of less developed countries, which cannot afford strict foreclosure policies like developed countries.

With the exception of the mortality rate parameter, the model does not have a strong link with health resilience citizens. The mortality rate parameter could also be related to external factors such as air pollution, which makes people more susceptible to respiratory diseases.

---

[5] Shi, P.; Cao, S.; Feng, P. SEIR Transmission dynamics model of 2019 nCoV coronavirus with considering the weak infectious ability and changes in latency duration.MedRxiv Infect. Dis. (Except HIV/AIDS) 2020.

[6] Shaikh, A.S.; Shaikh, I.N.; Nisar, K.S. A Mathematical Model of COVID-19 Using Fractional Derivative: Outbreak in India with Dynamics of Transmission and Control. Preprints 2020.

The introduction of Google's COVID-19 Community Mobility Report represents a restriction that was easily implemented in the model. New studies on the quality of these data and rigorous implementation could represent a new and interesting research topic.

The researchers thought that many of these questions could remain open, but the critical point of the study is not to determine exactly how each external factor influences the trend of infectious cases, since the researcher analyzed a problem for multiple facets from a point of view global. Furthermore, it was suspected that the official data and considerations were not sufficiently precise to serve as the basis for a very detailed study.

5. Conclusions

As a result of the conclusion of this research, the researchers applied two different approaches to solve the equations of the SEIR model to describe the evolution of the epidemic phenomenon in Italy and in the most affected regions of northern Italy (Lombardy, Veneto and Piedmont). They examined all possible data available as of April 15, 2020.

The main results indicate that the deterministic approach is not suitable for exploring possible solutions of the spatial domain because the mathematical problem is underdetermined. They recommended fitting the data for this outbreak using a stochastic approach, such as the PSO method. Taking advantage of the PSO approach, the researchers estimated different scenarios for a 30-day epidemic evolution. Each scenario refers to a different set of parameters estimated by the algorithm. The predicted scenarios are quite similar and suggest that each Italian region will reach the peak of the epidemic in mid-May. The influence of the time-varying infection rate on the prediction of the model may open interesting discussions about the effect of lockdown policies on the course of the epidemic in the near and distant future.

Since the model was delivered quickly and the study was carried out during the international emergency, they did not explore further the implications of the different "reopening" scenarios. Therefore, they could not predict the situation and the solution as a whole. Thus, they said that if the parameter remains at current values, for example, if the lockout policies are maintained or, better, the reopening is made with special attention to health security procedures, the prediction of the trend. of recovered and fatalities could be considered reliable, with the approximations and uncertainties highlighted by the PSO model. At the Italian level, despite the wide dispersion in the prediction of quarantined and recovered cases, the number of deaths will reach a number of around 33,000 to 35,000 cases by the end of May and the number of active cases will gradually decrease.

This prediction cannot consider the impact of future decisions on social distancing. The data and model predictions confirm that some valuable lessons must be learned from the approaches in South Korea, which was able to contain contagion very shortly before the infection spread widely. The Veneto region has been one of the best examples in Italy of how integrated and synergistic regional policies in social distance and the health system can cope with the epidemic, and its epidemiological scenario is now more optimistic than that. of Lombardy and Piedmont. The researchers noted that tailored actions provide much better epidemiological results than blanket blockades, also taking into account the example of South Korea.

The main objective of this work was to provide a new discussion and new tools capable of supporting policy makers in their decision to act to minimize the impact of the disease. Analysis

shows that because the Italian healthcare system is highly decentralized, different regions.

They managed different policies, which strongly influenced the evolution of the epidemic in its first months: the data and the prediction model well reflected the different approaches taken by Lombardy and Veneto, two regions with a similar socioeconomic fabric. The general lesson that could be drawn from this analysis goes beyond the mathematical model itself and will require a broader assessment of all possible socio-economic and political factors, even if analysis of the data from the situation in Veneto could be used to re- visit regional and central themes. policies from the beginning. If so, the regions will emulate the virtuous approach of Veneto, including more demanding requests to improve their diagnostic capacity that will weigh on the central government.

References:

Research Papers & Articles

- Godio, A.; Pace, F.; Vergnano, A. SEIR Modeling of the Italian Epidemic of SARS-CoV-2 Using Computational Swarm Intelligence. Int. J. Environ. Res. Public Health 2020, 17, 3535.
- Chatterjee K.; Chatterjee K.; Kumar A.; Shankar S. Healthcare impact of COVID-19 epidemic in India: A stochastic mathematical model. Medical Journal Armed Forces India, Elsevier, April 2020.
- Tayal, Devendra & Jain, Arti & Arora, Surbhi & Agarwal, Surbhi & Gupta, Tushar & Tyagi, Nikhil. (2014). Crime detection and criminal identification in India using data mining techniques. AI & SOCIETY. 30. 117-127. 10.1007/s00146-014-0539-6.
- Höhle, Michael &Jørgensen, Erik & O'Neill, Philip. (2005). Inference in disease transmission experiments by using stochastic epidemic models. Journal of the Royal Statistical Society: Series C (Applied Statistics). 54. 349 - 366. 10.1111/j.1467-9876.2005.00488.x.
- Webliography:
- http://rrcf.lagosirenetta.it/seir-model-r-code.html
- https://www.mdpi.com/1660-4601/17/10/3535#cite
- https://medworm.com/research/coronavirus.xml
- https://www.researchgate.net/publication/272031216_Crime_detection_and_criminal_identification_in_India_using_data_mining_techniques
- https://opengovasia.com/exclusive-opengov-breakfast-insight-session-on-analytics-leadership-in-government-transformation/
- https://hubpages.com/education/WhyTeachersMustUnderstandChildDevelopment
- https://www.15minutenews.com/business/2020/08/25/#!
- https://www.researchgate.net/publication/227689309_Inference_in_disease_transmission_experiments_by_using_stochastic_epidemic_models
- https://judithcurry.com/2020/05/10/why-herd-immunity-to-covid-19-is-reached-much-earlier-than-thought/
- https://www.medrxiv.org/content/10.1101/2020.06.26.20141077v2