



Comparing epidemiological models with the help of visualization dashboards

Csaba FARKAS

Sapientia Hungarian University of
Transylvania, Cluj-Napoca, Romania
Dept. of Mathematics and Informatics,
Târgu Mureş

email: farkascs@ms.sapientia.ro

David ICLANZAN

Sapientia Hungarian University of
Transylvania, Cluj-Napoca, Romania
Dept. of Mathematics and Informatics,
Târgu Mureş

email: iclanzan@ms.sapientia.ro

Boróka OLTEAN-PÉTER

Sapientia Hungarian University of
Transylvania, Cluj-Napoca, Romania
Dept. of Mathematics and Informatics,
Târgu Mureş

email:

boroka.oltean@ms.sapientia.ro

Géza VEKOV

Babeş-Bolyai University, Cluj-Napoca,
Romania
Faculty of Mathematics and Computer
Science

email: geza.vekov@cs.ubbcluj.ro

Abstract. In 2020, due to the COVID – 19 pandemic, various epidemiological models appeared in major studies [16, 22, 21], which differ in terms of complexity, type, etc. In accordance with the hypothesis, a complex model is more accurate and gives more reliable results than a simpler one because it takes into consideration more parameters.

In this paper we study three different epidemiological models: a SIR, a SEIR and a SEIR – type model. Our aim is to set up differential equation models, which rely on similar parameters, however, the systems of equation and number of parameters deviate from each other. A [visualization dashboard](#) is implemented through this study, and thus, we are

Computing Classification System 1998: G.2.2

Mathematics Subject Classification 2010: 68R15

Key words and phrases: Visualization dashboard, epidemiological model, optimization algorithm

able not only to study the models but also to make users understand the differences between the complexity of epidemiological models, and ultimately, to share a more specific overview about these that are defined by differential equations [24].

In order to validate our results, we make a comparison between the three models and the empirical data from Northern Italy and Wuhan, based on the infectious cases of COVID-19. To validate our results, we calculate the values of the parameters using the Least Square optimization algorithm.

1 Introduction

The COVID-19 pandemic has been responsible for over 24 million cases worldwide according to WHO reports. Not only has it been causing one of the biggest global health crises and the greatest challenge we have faced since World War II., but will also turn global economic growth “sharply negative” this year, based on a forecast by [BBC](#).

Mathematical models have been employed to inform media, authorities and researchers from different areas about the effect of the virus from different perspectives. In this study, we are going to discuss three different epidemiological models, which are visualized by a reactive tool. As [24] says epidemiological models are a key tool to guide public health measures. Without having experiences in crises such as this one, modelling and simulations require assumptions and different test scenarios. Therefore, visualizing three models with different complexity and parameter numbers as well as, comparing them are crucial to be well prepared for future events and to understand the roles of different factors in this pandemic.

Our first goal is to minimize the gap between medical reports, statistics, interoperability and public health information system [1, 6, 14, 3]. However, [3] study also highlighted that the pressure from this gap has always been particularly acute for the surveillance and management of infectious diseases with pandemic or bioterrorism potential, we reckon that nowadays this issue is one of the most pressing global problems.

We think that this tool is useful not only from the perspective of the public health information system, but also from that of simulations. We use one of the newest Javascript libraries, [Svelte](#), to create a reactive tool. Secondary, our aim is to validate our results with help of the Least Square optimization algorithm. Truncated Newton method is suitable for solving large nonlinear optimization problems [18].

The novelty of this research is that even though there exist several tools which visualize the results of mathematics models, which are presented in 3 section, we have no knowledge of tools which compare the models and validate the results with empirical data.

2 Short overview of epidemiological models and visualization dashboards

Study [3] in 2014 made a systematic review about visualization and analytics tools used for infectious disease epidemiology. In this study 247 articles were screened, and 88 articles were included in the review process. These articles primarily included descriptive reports, qualitative (e.g. interviews, focus groups) and usability studies. Although, public health workspace is extremely diverse [19] and the need for rapid access to information to support critical decisions in public health is inevitable, the public health information sources are unstandardized [12]. As a result, the visualization and analytics tools are various, especially in the case of COVID – 19 disease.

There are many studies which concentrate on the model proposal, such as [16, 22, 21], but there is a lack of visualization of these models. To our knowledge, the number of dashboards which project and simulate the population's exposure is very low, there is only one tool¹ which is a reactive data visualization based on an epidemiological model. However, the above mentioned visualization dashboard uses a basic SEIR model which cannot provide sufficiently accurate results from our point of view. We reckon that the demand for these tools which predict cases would be much higher, especially because there are a lot of different other software which visualize the empirical, measured cases in different countries and regions, such as Covidvisualizer², Gisanddata³ which was presented in [9] study or Wolframcloud Visualization Dashboard⁴. Other web pages which try to inform individuals and to minimize the COVID – 19 damages are rife, for instance, the 'plugandplaydiagnostics' software, presented in [25] study, which helps to prevent future epidemics or the COVID-19 Search Intensity Monitoring tool⁵. We also need to mention Epirisk dashboard⁶

¹<https://gabgoh.github.io/COVID/index.html>

²<https://www.covidvisualizer.com/>

³<https://gisanddata.maps.arcgis.com/apps/opsdashboard/index.html#/bda7594740fd40299423467b48e9ecf6>

⁴<https://www.wolframcloud.com/obj/examples/COVID19Dashboard>

⁵<https://covid19map.uptodate.com/>

⁶<https://epirisk.net/>

which is a computational platform estimating of the probability of exporting infected individuals from sites affected by a disease outbreak to other areas in the world through the airline transportation network and the daily commuting patterns.

As it can be seen, the above listed tools do not concentrate on mathematical models, and they do not contain a model comparison, even though different mathematical models bring various results [13, 1].

There is no accepted consensus regarding the modelling approach is considered to be the most accurate. Cooper et al. [8] use an susceptible-infected-removed (SIR) model and Wangping et al. [26] calculates with an extended SIR model, where transmission can be changed through many interventions, such as personal protective measures, community-level isolation, and city blockade. A lot of studies use susceptible-infected-exposed-removed (SEIR) models and their extended versions [29, 16, 22, 21, 23, 28, 17, 11].

In the above mentioned studies not only are the epidemiological models are different, but also the results and the parameter values are various.

3 Mathematical background for epidemiological models and optimization algorithm

Not only do we confirm that it is not possible to decide which model brings the most accurate results, but we also reckon that users' knowledge regarding the COVID – 19 pandemic and mathematical modelling is very wide-ranging. Even though we would accept the hypothesis according to which a more complex model is more accurate than a simpler one, we are faced with challenges regarding the user environment: users need to set up many parameters, which is difficult to comprehend. As one of our aims is to minimize the gap between general users and the public health information system, we reckon that by choosing a too complex model we would lose a significant part of possible users who we want to address.

For that, we have implemented three different models, with different complexity and different parameter numbers, but they use same parameters, which have the same meaning from a medical and environmental point of view.

3.1 SEIR-type model

The most complex model is a general SEIR-type model, which incorporates biological, social, environmental processes, such as governmental actions, (e.g.

school closing), weather conditions (temperature, humidity), and behavioral responses. Taking into account the above factors, we propose for the visualization the following SEIR-type model (see for more details see [10]):

$$\begin{cases} S' = -(\beta c(t) + c(t)q(1-\beta))S(I + \theta A)/N + \lambda S_q \\ E' = \beta c(t)(1-q)S(I + \theta A)/N - \sigma E \\ I' = \sigma \rho E - (\delta_I + \alpha + \gamma_I)I \\ A' = \sigma(1-\rho)E - \gamma_A A \\ S_q' = (1-\beta)c(t)qS(I + \theta A)/N - \lambda S_q \\ E_q' = \beta c(t)qS(I + \theta A)/N - \delta_q E_q \\ H' = \delta_I I + \delta_q E_q - (\alpha + \gamma_H)H \\ R' = (\delta_I + \alpha + \gamma_I)I + \gamma_A A + \gamma_H H - \gamma_R R \end{cases}$$

where the functions S, E, I, A, S_q, E_q, H and R denote the proportion of the population into eight groups: susceptible ($S(t)$), exposed ($E(t)$), infectious ($I(t)$), pre-symptomatic ($A(t)$), hospitalized ($H(t)$), recovered ($R(t)$), quarantined susceptible ($S_q(t)$) and isolated exposed ($E_q(t)$) groups of population, see Figure 1 for infection dynamics:

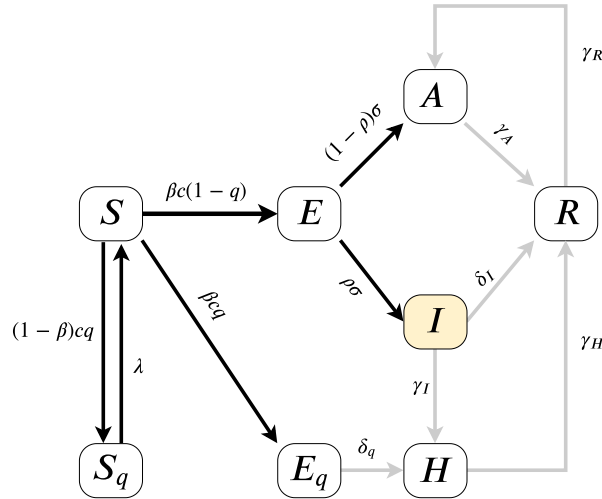


Figure 1: Model diagram for infection dynamics (SEIR-type model)

3.2 SEIR model

The SEIR – type model's simplified version is the SEIR model which uses significantly less parameters. This model was developed from the SEIR – type model:

$$\begin{cases} S' = -(\beta c(t) + c(t) q(1 - \beta)) SI/N \\ E' = (\beta c(t) + c(t) q(1 - \beta)) SI/N - \sigma \rho E \\ I' = \sigma \rho E - (\delta_I + \alpha + \gamma_I) I \\ R' = (\delta_I + \alpha + \gamma_I) I - \gamma_R R \end{cases}$$

where the functions denote the proportion of the population into four groups: susceptible ($S(t)$), exposed ($E(t)$), infectious ($I(t)$) and recovered ($R(t)$), see Figure 2.

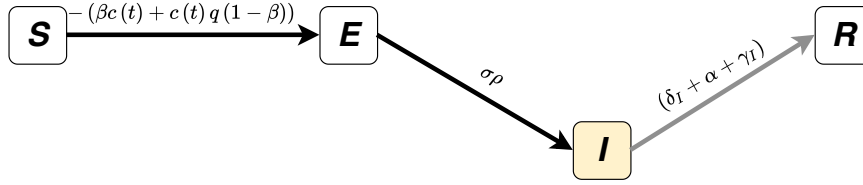


Figure 2: Model diagram for infection dynamics (SEIR model)

3.3 SIR model

By following the previous logic, we get the SIR model, if we exclude from the SEIR model the exposed population:

$$\begin{cases} S' = -\beta SI/N \\ I' = \beta SI/N - \gamma_I I \\ R' = \gamma_I I - \gamma_R R \end{cases}$$

where the population is split into 3 groups: susceptible ($S(t)$), infected ($I(t)$) and recovered ($R(t)$) proportions.

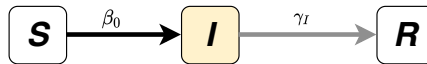


Figure 3: Model diagram for infection dynamics (SIR model)

3.4 The meaning of the parameters

The above mentioned models use same parameters in order to be comparable. These parameters are presented briefly, in case any additional information is needed, read [10] study.

The parameters σ and λ describe the transition rate of exposed individuals to the infected class and the rate at which the quarantined uninfected were released into the wider community, while the parameter ρ represents the probability of having symptoms among infected individuals. The parameters δ_I and δ_q denote the transition rate of symptomatic infected and quarantined exposed to the quarantined infected class. The γ_I , γ_A and γ_H represent the recovery rate of symptomatic, asymptomatic and quarantined infected individuals, and finally γ_R is the rate at which immunity is lost and recovered individuals move to the pre-symptomatic class (according to a recent NHK-World Japan report⁷ and [15]). The parameter θ represents the relative transmission probability of pre-symptomatic individuals to infected individuals. Finally we assume that natural birth and natural death rates are equal.

The motivation of such a choice is the following (see also [5, 28]): individuals move from quarantined cases with $1 - q$ proportion to S_q and with q proportion to E_q . If the transmission probability is β and the contact rate is c , then, the infected quarantined individuals move to E_q at rate of $\beta c q$ and uninfected quarantined individuals move to S_q at $(1 - \beta) c q$ rate. In case of not quarantined infected, they are going to move to E at a rate of $\beta c (1 - q)$. When an epidemiological outbreak occurs, many preemptive actions can be taken to mitigate the spreading. Once people become informed, they can change their behavior, working from home, practicing social distancing, and take actions like washing hands more often, wearing protective clothing, disinfecting etc., all of them contributing to the prevention of the spread. The media interacts with the susceptible population, it starts influencing them to take appropriate measures to minimize the chances of getting infected. This media influence is initially low and increases as the infection increases. This observation suggests the following contact rate function:

$$c(t) = c_a + \frac{3(c_0 - c_a)}{1 + 2b^{-t}}, \quad t \geq 0, \quad (1)$$

where $b < 1$ and c_0 denotes the initial contact rate, while c_a denotes the minimum contact rate under the current control strategies.

⁷https://www3.nhk.or.jp/nhkworld/en/news/20200315_13/

Obviously, not every parameter appears in all of three models. This description can be seen complete only in case of the SEIR – type model.

3.5 Least square optimization algorithm

Optimization algorithms can be classified into two major categories: line search methods and trust region methods, as Ya-xinag Yuan confirms [30]. The trust region approach associates with approximation, assuming that we have a current guess and the model can be constructed near that point. This algorithm is proposed to solve large-scale bound constrained minimization problems. It solves trust-region subproblems iteratively, augmenting with trust-region shape determined by the distance from the bounds and the direction of the gradient and by a special diagonal quadratic term as it is formulated in Python documentation⁸. The aim of this improvement is to iterate through the whole space of variables and to avoid hitting directly the bounds.

The mathematical approach of Trust Region Reflective Algorithm was formulated based on [“A subspace, interior and conjugate gradient method for large-scale bound-constrained minimization problems”] paper by Nikolay Mayorov through article⁹. It is defined the following bound-constrained minimization problem

$$\min f(x), \quad x \in \mathcal{F} = \{x : l \leq x \leq u\},$$

where $l \in \{\mathbb{R} \cup \{-\infty\}\}^n$ and $u \in \{\mathbb{R} \cup \{\infty\}\}^n$ and also the f function is a smooth function [2]. The $g(x) = \nabla f(x)$ and $H(x) = \nabla^2 f(x)$. Defining the following vector, we get

$$v(x)_i = \begin{cases} u_i - x_i & g_i < 0, u_i < \infty, \\ x_i - l_i & g_i > 0, l_i > -\infty, \\ 1 & \text{otherwise.} \end{cases}$$

Defining a matrix $D(x) = \text{diag}\left(v(x)^{\frac{1}{2}}\right)$, we can formulate the first order optimality as followed:

$$D^2(x) g(x) = 0.$$

If $v(x)_i = 0$ the Jacobian of the left hand does not exists, so we can consider that

$$v(x)_i \neq 0,$$

⁸https://docs.scipy.org/doc/scipy/reference/generated/scipy.optimize.least_squares.html

⁹<https://nmayorov.wordpress.com/2015/06/19/trust-region-reflective-algorithm/>

for all i . This happens if x is not on the bound. In this case the Newton step for this system is

$$\left(D^2 H + \text{diag}(g(x)) J_v \right) p = -D^2 g(x),$$

where J_v is the $v(x)$ vector's diagonal Jacobian matrix. Now the corresponding trust-region problem can be formulated:

$$\min_p m(p) = \frac{1}{2} p^T B p + g^T p, \text{ s.t. } \|D^{-1} p\| \leq \Delta,$$

where $B = H + D^{-1} C D^{-1}$ and $C = \text{diag}(g) J_v$. As it is formulated in [7] study reflective algorithms are used to maintain feasibility by a piecewise linear function, which helps to avoid bounds. The following implementation (Algorithm 3.5), which was developed by Nikolay Mayorov is mainly the same as `scipy.optimize.least_squares`:

```

1 import numpy as np
2
3 def reflective_transformation(y,l,u):
4     if l is None:
5         l=np.full_like(y, -np.inf)
6     if u is None:
7         u=np.full_like(y,np.inf)
8     l_fin=np.isfinite(l)
9     u_fin=np.isfinite(u)
10    x=y.copy()
11    m=l_fin & ~u_fin
12    x[m] = np.maximum(y[m], 2*l[m]-y[m])
13    m=~l_fin & u_fin
14    x[m] = np.maximum(y[m], 2*u[m]-y[m])
15
16    m=l_fin & u_fin
17    d=u-l
18    t=np.remainder(y[m]-l[m], 2*d[m])
19    x[m]=l[m]+np.minimum(t,2*d[m]-t)
20
21    return x
22

```

Algorithm 1: Reflective transformation algorithm

4 The overview of the visualization dashboard

Wekler et al. [27] defines dashboards as: “a visual display of data used to monitor conditions and/or facilitate understanding”, our aims are to monitor conditions and facilitate the understanding of the COVID – 19 pandemic and similar infectious diseases. The presented tool is a functional genre of dashboard which, as Sarikaya et al. [20] defines, means an interactive display that enables real-time monitoring of dynamically updating data.

The dashboard is a reactive, interactive tool, which is going to be presented through the following types of interactivity [20]:

1. Construction and Composition,
2. Multipage,
3. Interactive Interface.

The role of this classification is that reactivity and interactivity can take place at a number of different places in the dashboard lifecycle.

Reactive programming means a declarative programming paradigm where variables are updated automatically whenever other values change, while the re-execution of the statements is not necessary. Reactivity is programming with asynchronous data streams¹⁰. The benefit of the reactivity is that the dashboard becomes highly interactive with a multitude of UI events related to data events. This benefit evolves real-time monitoring: modifying a single value, such as a parameter, can automatically trigger other contents. Using Svelte Javascript library reactivity is realized by techniques such as virtual DOM which runs at build time, converting the components into a highly efficient imperative code that updates the DOM¹¹.

Due to the reactivity of the tool, the differential equation solver needs to compute the model very efficiently in order to serve a reasonable re-computation time based on the current parameters. From data visualization’s perspective one of the most pressing issues was to find an implementation which is accurate enough and does not cause lagging. For that, we have chosen the classic Runge-Kutta IV (RK4) method. There are a lot of Javascript libraries which implement Runge-Kutta methods, such as Runge-Kutta 4 library¹², or Cash-Karp implementation¹³ which is an adaptive Runge-Kutta method and

¹⁰<https://gist.github.com/staltz/868e7e9bc2a7b8c1f754>

¹¹<https://svelte.dev/blog/svelte-3-rethinking-reactivity>

¹²<https://www.npmjs.com/package/runge-kutta-4>

¹³<https://www.npmjs.com/package/ode45-cash-karp>

is presented in [4] study. However, we chose almost the simplest library Runge-Kutta library¹⁴, because this one can set up properly the interval and the step size. Our issue with adaptive Runge-Kutta method is that it gave significantly less efficiency than our final choice.

4.1 Construction and composition

These dashboards provide flexibility for the viewer to customize the placement of views, modify the visual representations inside those views, or select the particular dimensions and measures to visualize [20].

As we presented in Section 3, we visualize values of three different mathematical compartment models. As a result, users can modify the visual representations of the models. Users can choose the following three major representations:

- SIR model representation
- SEIR model representation
- SEIR — type model representation
- Overall trend.

As we discussed in section 1, different parameter lists help the user to find the model which is consistent with their knowledge. For instance, if the most complex model, 17 parameters and 8 initial values of functions appear (Figure 5). In case of selecting SEIR representation, the number of parameters is 11, 4 initial values are needed. For the least complex model only 5 parameters, 3 initial values are needed. These options modify drastically the user experience, because as long as the most complex representation gives a scientific view of the topic, the simplest representation can be understood by general users.

¹⁴<https://www.npmjs.com/package/runge-kutta>

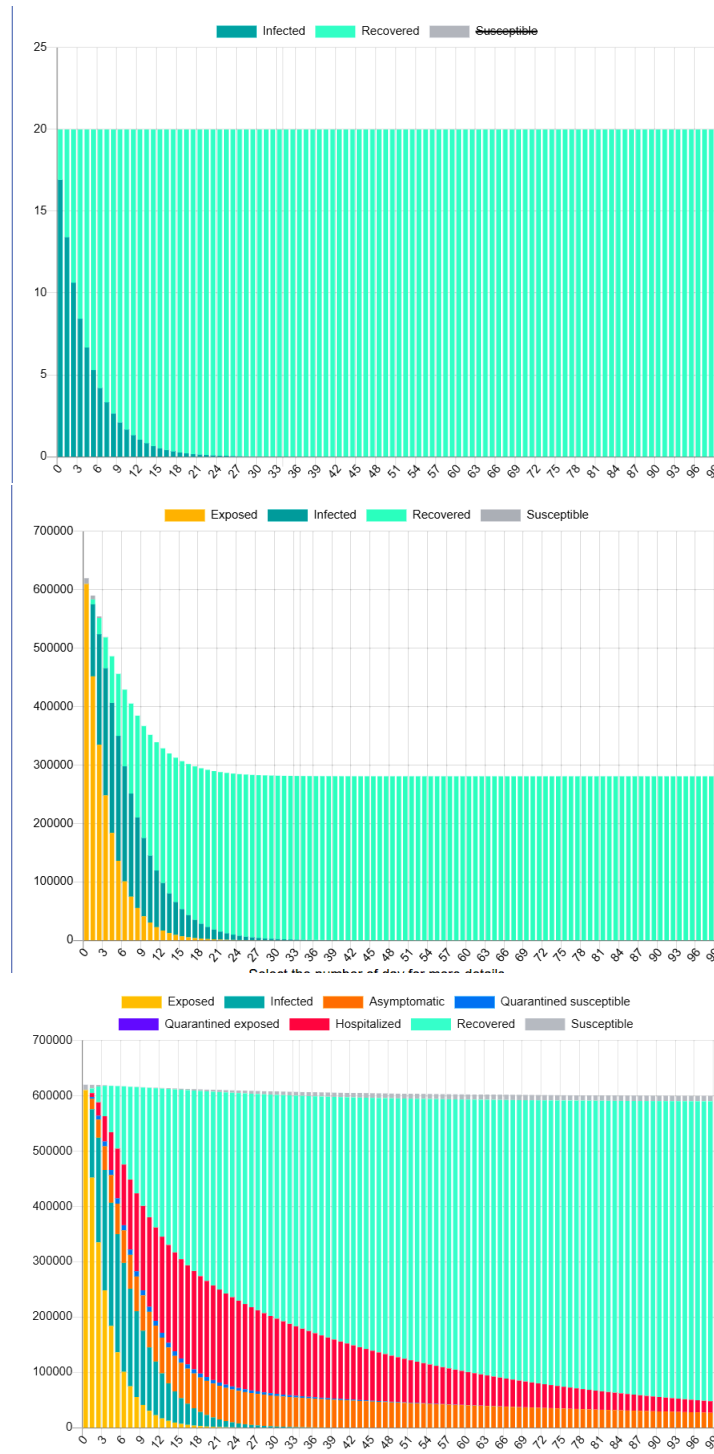


Figure 4: SIR model, SEIR model, SEIR-type models' representation, and overall trend

Definition:	Range of value:	Selected value:	Included in model:
c_0		41.85	<input checked="" type="checkbox"/>
c_a		1.287039726	<input checked="" type="checkbox"/>
q_1		1.484170130317585e-7	<input checked="" type="checkbox"/>
β_0		2.1031894440837713e-12	<input checked="" type="checkbox"/>
ϵ		0.88	<input checked="" type="checkbox"/>
σ		0.3333	<input checked="" type="checkbox"/>
λ		0.071428	<input checked="" type="checkbox"/>
δ_I		0.121	<input checked="" type="checkbox"/>
δ_q		0.03363	<input checked="" type="checkbox"/>
γ_I		0.1351146	<input checked="" type="checkbox"/>
γ_A		0.011795	<input checked="" type="checkbox"/>
γ_H		0.028635	<input checked="" type="checkbox"/>
γ_R		0.000001714578192	<input checked="" type="checkbox"/>
θ		0.00016357	<input checked="" type="checkbox"/>
α_b		0.0012209	<input checked="" type="checkbox"/>

Definition:	Range of value:	Selected value:
S		52740000
E		19.22
I		16.93
A		17.99
S_q		15.54
E_q		1.72
H		46.01
R		3.07

Figure 5: SEIR-plus model's parameter list and initial value condition

4.2 Multipage

Usually dashboards are monopages, but some of them support tabbed layouts. These dashboards allow viewers to switch between pages, which may have visualizations that relate to a different component of decision-making or help to provide the necessary context [20].

Regarding the presented dashboard, multipage and constraction-composition relate closely. The user can navigate between the three model representations and the comparison of three models. However, the dashboard has multi-level structure, meaning that a single page, such as Overall trend page, is divided to more than one components. The user can compare the tree models based on each basic proportion of population. As it can seen on 6, the comparison of models is visualized based on Recovered ($R(t)$) cases.

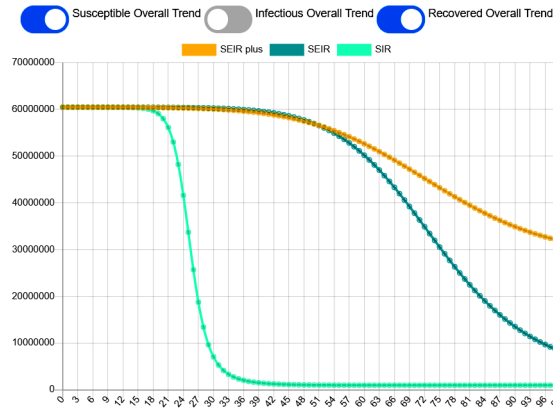


Figure 6: The user can choose which function to be visualized: $S(t)$, $I(t)$ or $R(t)$. Here, the $S(t)$ function is displayed.

4.3 Interactive interface

Obviously, drop-down menus (Figure 7), slicers (Figure 5) appear on the dashboard which improve user-experience. But, due to the reactivity of the tool, the differential equation solver was needed the model to be computed very efficiently in order to serve a reasonable re-computation time based on the current parameters. From data visualization one of the most pressing issues was to find an implementation which is accurate enough and does not cause lagging. For that, we have chosen the classic Runge-Kutta IV (RK4) method. There are a lot of Javascript libraries which implement Runge-Kutta meth-

ods, such as Runge-Kutta 4 library¹⁵, or Cash-Karp implementation¹⁶ which is an adaptive Runge-Kutta method and is presented in [4] study. However, we chose one of the simplest libraries Runge Kutta library¹⁷, because the adaptive Runge-Kutta method gave significantly less efficiency than our final choice.

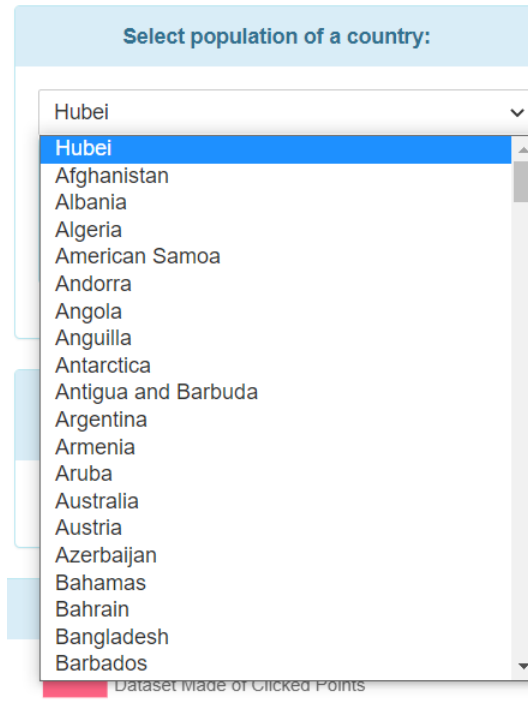


Figure 7: Drop-down list where the user can choose the country or area data the data of which they want to be set up.

5 Results

We reckon that we accomplished our aims formulated in Section 1. Our primarily one with visualization dashboard was to reduce the gap between general user's knowledge regarding mathematical modelling and public health information system which helps to make correct decisions. For that, we implemented different informative labelling and description (as shown Figure 8) regarding

¹⁵<https://www.npmjs.com/package/runge-kutta-4>

¹⁶<https://www.npmjs.com/package/ode45-cash-karp>

¹⁷<https://www.npmjs.com/package/runge-kutta>

the models, the parameters and initial conditions, which change dynamically after modifying of the models. The fact, that parameters can be modified only in predefined intervals serves the same purpose. We would like to suggest values which reflect the reality. Because of it, when the dashboard is loaded for the first time, the parameters are predefined based on Least-Square optimization algorithm and empirical data of Northern Italy which can be seen in SEIR-type column of Table 2.

Information about model	
<p>Differential equation system:</p> $S' = -[\beta(t)c(t) + c(t)q(t)(1 - \beta(t))]S(I + \theta A) + \lambda S_q$ $E' = \beta(t)c(t)(1 - q(t))S(I + \theta A) - \sigma E$ $I' = \sigma E - (\delta_I + \alpha + \gamma_I)I$ $A' = \sigma(1 - \varrho)E - \gamma_A A + \gamma_R R$ $S_q' = (1 - \beta(t))c(t)q(t)S(I + \theta A) - \lambda S_q$ $E_q' = \beta(t)c(t)q(t)S(I + \theta A) - S_q E_q$ $H' = \delta_I I + \delta_q E_q - (\alpha + \gamma_H)H$ $R' = \gamma_I I + \gamma_A A + \gamma_H H - \gamma_R R$	<p>Other functions in the system:</p> $c(t) = c_a + 3(c_0 - c_a)/(1 + 2b^{-t})$ $q(t) = (q_1 t + q_0)/(t + 1)$ <p>In case of the presence of humidity and temprature:</p> $\beta(t) = (1 - \alpha_\beta)\beta_0 \cdot (1 + \xi AH(t)) \cdot (1 - (I(t) + A(t))/(S(t) + R(t)))^2$ <p>In case of the presence of temprature:</p> $\beta(t) = (1 - \alpha_\beta)\beta_0 \cdot (1 + \xi AH(t)) \cdot (1 - (I(t) + A(t))/(S(t) + R(t)))^2.$

Figure 8: Model description and informative labelling

However, based on our results, we reckon that choosing the proper model and parameters' values presumes prior knowledge from users. For that, we think all visualization dashboards based on mathematical models mostly targeted users who are mathematicians investigating the discussed models, and these tools might be providing them with important insights.

5.1 Visualization dashboard

The functionalities of the visualization dashboard's presented in this study accomplish the aims formulated in 1 Section. The following functionalities serve the demands of general lay users but they also help the work of practised ones by the following components.

As it is presented on Figure 9, the web page has different components. When the user enters the web page, they see a predefined parameter settings, ebpagehich we calculated based on the cases of infection as measured in the region of Northern Italy and the parameters were obtained with the help of the

Least – Square algorithm. The user can start to change parameters manually or can import new parameter setup via a JSON file. It is given the opportunity to set up some parameters manually after importing the parameters, and the other way around, they can import some parameters after setting up the parameter manually as well. After setting up the parameters, they can be exported, resulting a JSON file with the current parameter setup, or can directly display data visualization. If the user changes a parameter, the differential equation solver automatically recalculates the result. After data visualization, the charts can be exported in the PDF format as well.

Due to reactivity, these steps are not detached strictly, they can be inverted, and other steps will respond to these changes directly and instantly.

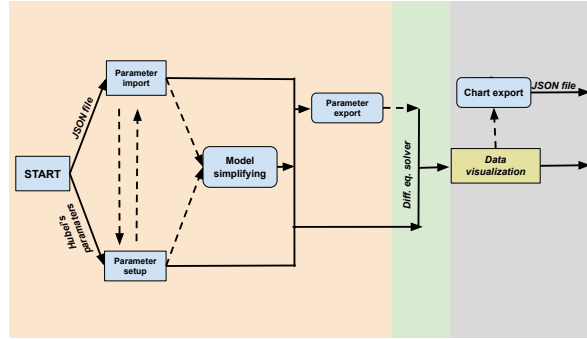


Figure 9: The structure of visualization dashboard, the yellow segment represents the parameter setup part, the green component represents the differential equation solver and the grey one represents data visualization.

5.2 Models comparison

As we expected, the tree models usually bring slightly different results (Figures 10, 11). These figures present results with different parameter setup and initial value conditions.

We can observe that the SIR epidemiological model presented in Section 3 is not complex enough to model COVID – 19 pandemic.

The comparison of models can be performed from various perspectives. We have studied the models based on the comparison of empirical data. For that, we have used Least-Square optimization algorithm to estimate the parameters for models. To see more about this optimization algorithm, see [10].

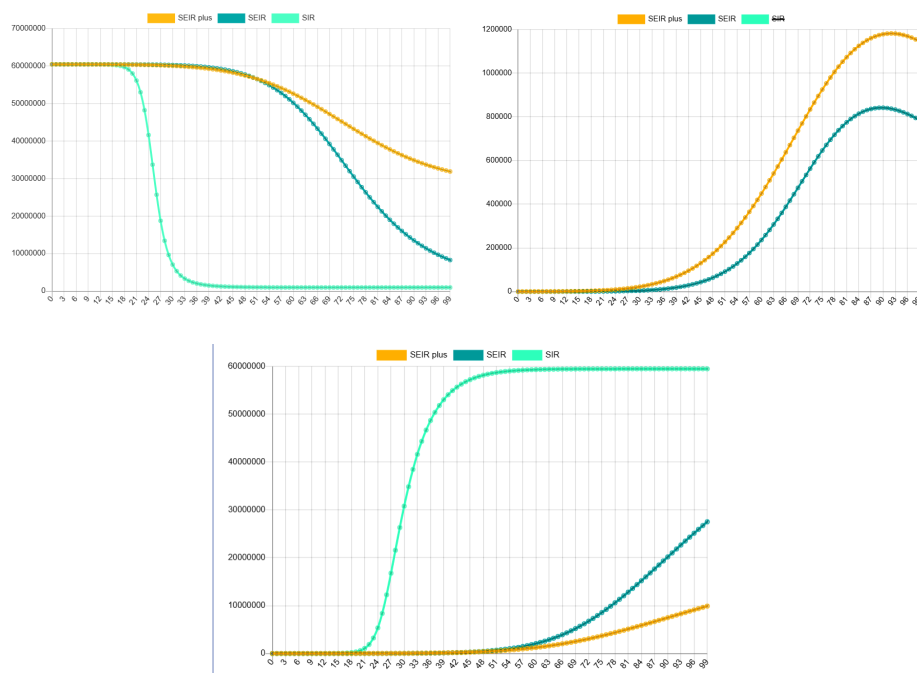
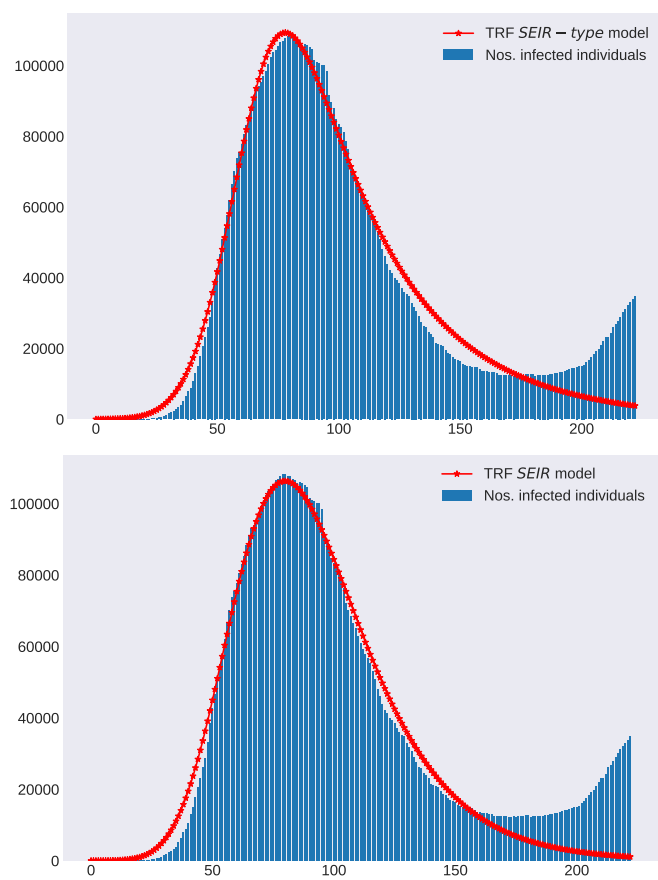


Figure 10: Model comparison (first fig. $S(t)$, second fig. $I(t)$, third fig. $R(t)$)



As long as the SEIR – type and SEIR model perform quite similarly, the SIR model fails from different perspectives. As shown in Table 1, the cumulative errors between measured and estimated values are similar in case of the SEIR – type model and the SEIR model. The algorithm was run for multiple dataset, and the errors move approximately in the same interval in case of SEIR and SEIR – type models. The algorithm could not fit to measured data in case of the SIR model, as the error values suggest.

Empirical data	SIR	SEIR	SEIR-type
North-Italy	160968187	1209142	1113692
Hubei	149720737	103984	127389
Germany	215503664	1446547	1336592

Table 1: Models comparison based on error bounds

As long as the error bounds of the SEIR and the SEIR-type models are approximately similar, the parameter lists based on the estimation are various. And even tough, the estimations seem to be very close to each other, as shown Figures 11, these results were gotten from strongly dissimilar parameter lists (see Table 2).

6 Discussion

The phenomenon, which was presented at Subsection 5.2 can have multiple reasons. We need to emphasize that running an optimization algorithm does not necessarily mean that parameters from Table 2 are the most optimal parameters for the presented models.

These approaches can be further aims, but our primarily goal with this study was not parameter optimization. In accordance with the previously formulated hypothesis, we cannot decide unequivocally that a more complex model is more accurate than a simpler one. This study highlights the fact that neither the complexity of the model, nor parameter number are crucial in mathematical models. If we do not interpret the values of parameters from medical and epidemiological point of view, almost every model can fit properly to empirical values.

We also need to emphasize that these models divide the population in different groups. For instance, as long as the SEIR-type model defines 8 different proportions, associating exposed, isolated exposed, etc., the SIR model does not even take into account the exposed proportion of the population. This dif-

ference obviously comes up in the values and in the visualization as well. Based on this study we confirm that mathematical epidemiological models provide a certain way to understand infectious diseases and pandemics, but without a medical perspective it is not possible to conclude clear conclusions regarding the future-events.

Based on the above formulated affirmations, we think that this visualization dashboard is the most useful, when the user knows some parameter values and wants to check different test scenarios based on their knowledge, and not inversely.

Name of variables	SEIR-type	SEIR
S	60461744	60461744
E	194	195.656
I	10	10
A	96.69	—
S_q	151.86	—
E_q	158.28	—
H	2	—
R	0	0
c_0	33.74	33.34
c_a	10.95	2.49
q_1	0.29	0.20
β_0	0.11	0.38
ϵ	0.46	0.57
σ	0.17	0.17
λ	0.071	0.071
δ_I	0.0028	0.999
δ_q	0.16	—
γ_I	0.219	0.333
γ_R	0.219	0.023
γ_A	0.197	—
γ_H	0.326	—
θ	0.502	—
α	0.838	0.999

Table 2: Optimal parameter lists

Acknowledgment

Csaba Farkas has been supported by the Sapientia Foundation — Institute for Scientific Research, Romania, Project No. 17/11.06.2019. Boróka Oltean-Péter has been supported by the Sapientia Hungariae Foundation – Collegium Talentum project and by Accenture Student Research Scholarship.

References

- [1] A. Abta, A. Kaddar, T. Hamad, Global stability for delay sir and seir epidemic models with saturated incidence rates, *Electronic Journal of Differential Equations* **2012**, 23 (2012) 1–13. [⇒ 261, 263](#)
- [2] M. Branch, Th. Coleman, Y. Li, A subspace, interior, and conjugate gradient method for large-scale bound-constrained minimization problems. *SIAM Journal on Scientific Computing*, 21, 1 (1999) 1–13. [doi:10.1137/S1064827595289108](#) [⇒ 267](#)
- [3] L. N. Carroll, A. P. Au, L. T. Detwiler, T. Ch. Fu, I. S. Painter, N. F. Abernethy, Visualization and analytics tools for infectious disease epidemiology: A systematic review, *Journal of Biomedical Informatics*, **51** (2014) 287–298. [doi:10.1016/j.jbi.2014.04.006](#) [⇒ 261, 262](#)
- [4] J. Cash, A. Karp, A variable order Runge-Kutta method for value problems with rapidly varying right-hand sides, *ACM Trans. Math. Softw.*, **16**, 3 (1990) 201–222. [doi:10.1145/79505.79507](#) [⇒ 270, 274](#)
- [5] C. Castillo-Chavez, C. W. Castillo-Garsow, A.-A. Yakubu, Mathematical models of isolation and quarantine, *JAMA*, 290, 21 (2003) 2876–2877. [doi:10.1001/jama.290.21.2876](#) [⇒ 266](#)
- [6] H. Chen, D. Zeng, P. Yan, Data visualization, information dissemination, and alerting, In *Integrated Series in Information Systems*, vol. 21. Springer, New York, NY. 2010, pp. 73–87. [doi:10.1007/978-1-4419-1278-7_5](#) [⇒ 261](#)
- [7] Th. Coleman Y Li, On the convergence of reflective newton methods for large-scale nonlinear minimization subject to bounds, *Math. Program.* **67**, 1-3 (1994) 189–224. [doi: 10.1007/BF01582221](#) [⇒ 268](#)
- [8] I. Cooper, A. Mondal, Ch. G. Antonopoulos, A SIR model assumption for the spread of COVID-19 in different communities, *Chaos, Solitons & Fractals*, **139** (2020) 110057. [doi:10.1016/j.chaos.2020.110057](#) [⇒ 263](#)
- [9] L. Gardner E. Dong, H. Du, An interactive web-based dashboard to track Covid-19 real time, *THE LANCET Infectious Diseases* **20**, 5 (2020) 533–534. [doi:10.1016/S1473-3099\(20\)30120-1](#) [⇒ 262](#)
- [10] Cs. Farkas, D. Iclanzan, B. Olteán Péter, G. Vekov, Estimation of parameters for a temperature and humidity-dependent compartmental model of the Covid-19 outbreak, *Preprint*, 2020. [⇒ 264, 266, 276](#)

- [11] Q. Griette, Z. Liu, P. Magal, Estimating the last day for Covid-19 outbreak in mainland China, *Preprint at medRxiv.org*, July 6, 2020.
[doi:10.1101/2020.04.14.20064824](https://doi.org/10.1101/2020.04.14.20064824) ⇒ 263
- [12] B. L. Humphreys, Meeting information needs in health policy and public health: Priorities for the national library of medicine and the national network of libraries of medicine, *Journal of Urban Health*, 75, 4 (1998) 878–883.
[doi:10.1007/BF02344515](https://doi.org/10.1007/BF02344515) ⇒ 262
- [13] A. Kaddar, A. Abta, H. T. Alaoui, A comparison of delayed SIR and SEIR epidemic models, *Nonlinear Analysis: Modelling and Control* **16**, 2 (2011) 181–190. [doi:10.15388/na.16.2.14104](https://doi.org/10.15388/na.16.2.14104) ⇒ 263
- [14] M. Klompas, M. Murphy, J. Lankiewicz, J. McVetta, R. Lazarus, E. Eggleston, P. Daly, P. Oppedisano, B. Beagan, Ch. Kirby, R. Platt, Harnessing electronic health records for public health surveillance, *Online Journal of Public Health Informatics*, **3**, 3 (2011) [doi:10.5210/ojphi.v3i3.3794](https://doi.org/10.5210/ojphi.v3i3.3794) ⇒ 261
- [15] G. Li, Y. Fan, Y. Lai, T. Han, Z. Li, P. Zhou, P. Pan, W. Wang, D. Hu, X. Liu, Q. Zhang, J. Wu, Coronavirus infections and immune responses, *Journal of Medical Virology*, **92**, 4 (2020) 424–432. [doi:10.1002/jmv.25685](https://doi.org/10.1002/jmv.25685) ⇒ 266
- [16] Q. Lin, S. Zhao, D. Gao, Y. Lou, S. Yang, S. S. Musa, M. H. Wang, Y. Cai, W. Wang, L. Yang, D. He, A conceptual model for the coronavirus disease 2019 (covid-19) outbreak in wuhan, china with individual reaction and governmental action, *International Journal of Infectious Diseases*, **93** (2020) 211–216. [doi:10.1016/j.ijid.2020.02.058](https://doi.org/10.1016/j.ijid.2020.02.058) ⇒ 260, 262, 263
- [17] Y. Ma, Y. Zhao, J. Liu, X. He, B. Wang, Sh. Fu, J. Yan, J. Niu, J. Zhou, B. Luo, Effects of temperature variation and humidity on the death of Covid-19 in Wuhan, China, *Science of the Total Environment*, **724**, 7 (2020) 138226. [doi:10.1016/j.scitotenv.2020.138226](https://doi.org/10.1016/j.scitotenv.2020.138226) ⇒ 263
- [18] S. G. Nash, A survey of truncated-Newton methods, *Journal of Computational and Applied Mathematics* **124**, 1-2 (2000) 45–59. [doi:10.1016/S0377-0427\(00\)00426-X](https://doi.org/10.1016/S0377-0427(00)00426-X) ⇒ 261
- [19] D. Revere, A. M. Turner, A. Madhavan, N. Rambo, P. F. Bugni, A. Kimball, Sh. S. Fuller, Understanding the information needs of public health practitioners: A literature review to inform design of an interactive digital knowledge management system, *Journal of Biomedical Informatics*, **40**, 4 (2007) 410–421. [doi:10.1016/j.jbi.2006.12.008](https://doi.org/10.1016/j.jbi.2006.12.008) ⇒ 262
- [20] A. Sarikaya, M. Correll, L. Bartram, M. Tory, D. Fisher, What do we talk about when we talk about dashboards? *IEEE Transactions on Visualization and Computer Graphics*, **29**, 1 (2019) 682–692. ⇒ 269, 270, 273
- [21] B. Tang, N. L. Bragazzi, Q. Li, S. Tang, Y. Xiao, J. Wu, The effectiveness of quarantine and isolation determine the trend of Covid-19 epidemics in the final phase of current outbreak in China, *International Journal of Infectious Diseases* **95**, 6 (2020) 288–293. [doi:10.1016/j.ijid.2020.03.018](https://doi.org/10.1016/j.ijid.2020.03.018) ⇒ 260, 262, 263
- [22] B. Tang, N. L. Bragazzi, Q. Li, S. Tang, Y. Xiao, J. Wu, An updated estimation of the risk of transmission of the novel coronavirus (2019-nCov), *Infectious Disease Modelling*, **5** (2020) 248–255. [doi:10.1016/j.idm.2020.02.001](https://doi.org/10.1016/j.idm.2020.02.001) ⇒ 260, 262, 263

- [23] B. Tang, X. Wang, Q. Li, N. L. Bragazzi, S. Tang, Y. Xiao, J. Wu, Estimation of the transmission risk of the 2019-nCov and its implication for public health interventions, *Journal of Clinical Medicine* **9**, 2 (2020) 462. [doi:10.3390/jcm9020462](https://doi.org/10.3390/jcm9020462) ⇒ 263
- [24] R. N. Thompson, Epidemiological models are important tools for guiding COVID-19 interventions, *BMC Medicine* **18**, 152 (2020). [doi:10.1186/s12916-020-01628-4](https://doi.org/10.1186/s12916-020-01628-4) ⇒ 261
- [25] B. Udugama, P. Kadhiresan, H. N. Kozlowski, A. Malekjahani, M. Osborne, V. Y. C. Li, H. Chen, J. B. Gubbay S. Mubareka, W. C. W. Chan, Diagnosing Covid-19: The disease and tools for detection, *ACS Nano* **14**, 4 (2020) 3822–3835. [doi:doi:10.1021/acsnano.0c02624](https://doi.org/10.1021/acsnano.0c02624) ⇒ 262
- [26] J. Wangping, H. Ke, S. Yang, C. Wenzhe, W. Shengshu, Y. Shanshan, W. Jianwei, K. Fuyin, T. Penggang, L. Jing, L. Miao, H. Yao, Extended SIR prediction of the epidemics trend of COVID-19 in Italy and compared with Hunan, China. *Frontiers in Medicine*, **7**, 5 (2020) [doi:10.3389/fmed.2020.00169](https://doi.org/10.3389/fmed.2020.00169) ⇒ 263
- [27] S. Wexler, J. Shaffer, A. Cotgreave, *The Big Book of Dashboards: Visualizing Your Data Using Real-World Business Scenarios*, Wiley Publishing, 1st edition, 2017. ISBN: 978-1-119-28271-6 ⇒ 269
- [28] Y. Xiao, S. Tang, J. Wu, Media impact switching surface during an infectious disease outbreak, *Scientific Reports* **5**, 7838 (2015). [doi:10.1038/srep07838](https://doi.org/10.1038/srep07838) ⇒ 263, 266
- [29] Z. Yang, Zh. Zeng, K. Wang, S.-S. Wong, W. Liang, M. Zanin, P. Liu, X. Cao, Zh. Gao, Zh. Mai, J. Liang, X. Liu, Sh. Li, Y. Li, F. Ye, W. Guan, Y. Yang, F. Li, Sh. Luo, Y. Xie, B. Liu, Zh. Wang, Sh. Zhang, Y. Wang, N. Zhong, J. He, Modified SEIR and AI prediction of the epidemics trend of COVID-19 in China under public health interventions, *Journal of Thoracic Disease* **12**, 3 (2020) 165–174. [doi:10.21037/jtd.2020.02.64](https://doi.org/10.21037/jtd.2020.02.64) ⇒ 263
- [30] Ya-xiang Yuan. A review of [trust region algorithms](#) for optimization, *ICM99: Proceedings of the Fourth International Congress on Industrial and Applied Mathematics*, September, 1999. ⇒ 267

Received: October 13, 2020 • Revised: November 11, 2020