A Contribution to Modeling and Computer Simulation of Species Spread in Natural Environments

Muris Torlak, Vahidin Hadžiabdić, and Sadjit Metović

Abstract In this paper, analogy of heat- or mass-transfer problem and progression of species in natural environments is used. Time-dependent spatial distributions of species in a given population (such as microbes, bacteria, viruses or here-with infected individuals etc.) are described by time-dependent convection-diffusion equation, which is widely used in analysis of a number of physical, engineering or environmental problems. In the presented example, a finite-volume method is used for its numerical solution. The source term is calibrated using the available recorded data on the species concentration development in the beginning stage. The integral results of the proposed model are compared with another, integral model (SIR) based on ordinary differential equations. Calculated spatial distributions for three different modeling cases are given.

Keywords Modeling · Computer simulation · Species · Partial differential equations · Disease spreading

1 Introduction

Analysis of dynamics of biological species spread in natural environments requires a variety of mathematical tools. Typically, these tools involve model creation, solving the corresponding equations, usually including or derived from ordinary or partial differential equations, statistical analysis and others.

A considerable number of models used for this purpose is described by ordinary differential equations, usually delivering integral analysis of the variations in time, while the spatial distributions remain unresolved. The spatial models of biological species in natural environments due to growth or flow-induced transport in aquatic ecosystems, e.g. plankton populations in the ocean under complex turbulent flows,
can be described by partial differential equations, such as used for heat or mass transfer, transport of species in fluids, semiconductor physics and other problems. The reaction–diffusion–advection equation is also used to model some of the problems, e.g. bacterial chemotaxis or migration of population.

On the other hand, simulations of some spreading phenomena, such as epidemics, usually consider variation of the sought variables (e.g. number of infected people) in time. There is a number of different models used for this (SIR, SIRS, SIS, SEIR, MSEIR etc.) [1–3]. The problems are analyzed for a given part of space, e.g. for a country or for a city, but without detailed information about spatial variation of the characteristic quantities and their interaction.

In this paper solution of the time-dependent convection–diffusion equation applied to spread of infectious disease, observing temporal and spatial variation of the number of infected individuals is tested. The problem is solved first using one of the frequently employed models, SIR model. Then the analogy between the heat-transfer problem and the disease spread is outlined, and finally numerical solution of the convection–diffusion model for disease spread is demonstrated.

2 Modeling by Ordinary Differential Equations

In a large number of cases, spread of infectious diseases is described by a system of non-linear, mutually coupled differential equations. The basic model of this kind, which is probably mostly used, is the SIR (Susceptible-Infected-Recovered) model [1].

The model emanates from the following assumptions: a closed population under consideration consists of three groups of individuals: susceptible, infected and recovered ones (S, I, R, respectively), and its total number is constant; a recovered individual cannot be re-infected; the rate of transmission of the disease is proportional to the number of possible communications between susceptible S and infected I individuals; the recovery rate is proportional to the number of infected. The unknown quantities are the number of susceptible, infected and recovered individuals which are variable in time. The model is mathematically described as follows:

\[
\begin{align*}
\frac{dS}{dt} &= -\beta SI, \\
\frac{dI}{dt} &= \beta SI - \nu I, \\
\frac{dR}{dt} &= \nu I,
\end{align*}
\]  

(1)

where \(\beta\) and \(\nu\) are positive parameters. The parameter \(\beta\) describes the transmission rate. The parameter \(\nu\) describes the recovery rate and it is the reciprocal value of the
average recovery time. The model described by Eq. (1) is suitable for simulations of relatively fast disease development, so that assumption on the constant total number of individuals in the observed time is valid. Beside this one, there are many other models in use [2, 3] with different degrees of complexity introduced to provide appropriate accuracy and reliability of predictions in different disease scenarios.

The model given by Eq. (1), as well as other similar models from the same class, describe the considered population in integral way: only variations in time are observed, and they do not contain any information about distribution of the quantities of interest in space. This lack of spatial variations is one of the drawbacks of the models based on time-dependent ordinary differential equations.

3 Convection–Diffusion Equation

In this work time-dependent convection–diffusion equation is used for analysis of infectious-disease development both in time and space. This type of equation is very frequently used in analysis of a number of physical and engineering problems, such as heat transfer or transport of species in a continuous medium. Very robust and efficient numerical solution methods are available since decades, delivering a good framework for application to analysis of temporal variation and spatial distribution of species in other systems with the same or similar behavior. One of such application fields could be spread of viruses which cause infectious diseases in a population, where the number of infected persons would be the primary quantity of interest. The equation is written here in integral form:

$$\frac{\partial}{\partial t} \int_V \rho c \phi dV + \int_S \rho \vec{v} c \phi d\vec{S} = \int_S k \vec{\nabla} \phi d\vec{S} + \int_V q^+ dV - \int_V q^- dV.$$  \hspace{1cm} (2)

Equation (2) describes temporal and spatial variation of the unknown quantity $\phi$ within a part of space, which is bounded by a closed surface $S$ and whose size is expressed by the volume $V$. The first term describes the local rate of change of the unknown quantity $\phi$. The second term describes convective transport through the boundary $S$, which may be applied to directed transport of the individuals at relatively large distances (in the further text also denoted as “convection”). The third term is the diffusion term, which is here used to describe “undirected” transport and spread of the quantity $\phi$ in the intermediate vicinity of the considered part of space (in the further text also denoted as “diffusion”). The last two terms describe local production and destruction rate, or simply local sources and sinks of the considered quantity. In case of modeling of chemical processes they describe the reaction effects. The quantities $\rho$, $c$, and $k$ are the properties describing behavior and response of the specific medium in the considered part of space. In heat transfer problem, they describe the density, the specific heat capacity, and the thermal conductivity of the material, respectively.
Table 1  Analogy between the quantities in heat transfer and disease spread

| Variable | Heat transfer                      | Disease spread                      |
|----------|-----------------------------------|-------------------------------------|
| \( \phi \) | Temperature                       | Specific number of infected persons\(^a\) |
| \( V \) | Volume                            | Populated area                     |
| \( \vec{S} \) | Surface/area                      | Border/width\(^b\)                 |
| \( \rho \) | Density                           | Population density                 |
| \( c \) | Specific heat capacity            | Resistance to disease               |
| \( \rho \vec{v} \vec{S} \) | Mass flux of the transporting medium | Transportation\(^c\)               |
| \( k \) | Thermal conductivity              | Local transmission capability       |
| \( q^+ \) | Specific heat source rate         | Production rate                     |
| \( q^- \) | Specific heat sink rate           | Destruction rate                    |

\(^a\) Fraction of the total population or concentration of the infected persons

\(^b\) Note that \( \vec{S} \) may be described in different ways. It may also be a number of neighboring "points", such as contact persons or neighboring places/towns

\(^c\) This quantity describes the net number of persons moving from one place to the other

Assuming analogy between the convective-diffusive transport and disease spreading, as described in Table 1, Eq. (2) can be used to simulate the temporal and spatial development of the disease.

Clearly, the results depend on the input data. While the geographic data (area, border length, population density) can be estimated with sufficient accuracy, the values such as production or destruction rate can be estimated only roughly, while the resistance to disease, the transportation, and the local transmission rate must be assumed since the detailed data on these effects are in a large number of cases not available.

When Eq. (2) is written in differential form (as a partial differential equation), finite-difference methods can be used for its solution. This is, however, difficult to process on complex geometries, so finite-element or finite-volume methods are usually preferred choice. In this work a finite-volume method, such as described in [4], is implemented in own computer program [5] and used for solution of Eq. (2). The procedure is here described just briefly, and the details can be found in the aforementioned literature.

The considered part of space is divided into a set of cells of polyhedral shapes, typically hexahedra or tetrahedra. In plane two-dimensional domains, they reduce to polygons, quadrilaterals and triangles, respectively; the volume integrals reduce to the integrals over the polygon area, while the surface integrals reduce to the integrals along their boundaries. The cells do not penetrate into each other and there are no voids between them. For each cell in the set Eq. (2) is discretized, yielding
thus a system of non-linear algebraic equations to be solved for the quantity $\phi$ at the computational points. In the method applied here, the computational points are located at the cell centroids. The discretization include approximation of the volume and the surface integrals by midpoint rule and assumption of linear variation of the variables in space. Consequently, the variable values required on the surface $\vec{S}$ are obtained by linear interpolation, except in case of strong convection where other techniques might be required (such as upwind schemes) in order to promote stability. Similarly, the required gradients at the surface $\vec{S}$ are approximated by central differencing scheme, while the gradients at the computational points are obtained by least-square-method in accordance with the assumed linear spatial variation of the variables. The time derivative is approximated by a backward differencing scheme, typically a 1st- or 2nd-order accurate one. The solution process is iterative, allowing temporary linearization of the equations in the system as well as explicit treatment of the numerical correction terms. Due to suitable matrix structure of the linearized system, in each iteration systems are solved by one of the methods for unstructured: preconditioned BiCGSTAB for non-symmetric matrices when the convection is included or preconditioned CG method for symmetric matrices. The iterations are repeated where the algebraic systems are recalculated, assembled and solved in turn, until prescribed tolerance level is reached.

4 Spread of COVID-19 Disease in Bosnia-Herzegovina

The models described by Eqs. (1) and (2) are applied to simulation of COVID-19 disease development in Bosnia-Herzegovina in the spring 2020. The total population in Bosnia-Herzegovina is 3.8 million people.

The parameters adopted for simulation using SIR model, Eq. (1), are explained as follows. Assuming that one infected individual infects the others every three days, the transmission rate $\beta = 1/3$ is used. The recovery rate $\upsilon$ is described as reciprocal value of the required recovery time. Typically, recovery takes from 14 to 28 days. Here the value of the recovery rate of $1/25 = 0.04$ is assumed. Since this is an initial-value problem, as described by Eq. (1), the results depend strongly on the initial values of the susceptible and the infected individuals. Their precise estimation is frequently difficult, not rarely it is impossible, and usually these values are not known in advance. In the present simulation we use a guess with 1 initial infected person and the entire population being susceptible, i.e. 3.8 million people. The time step size is 1 day. Explicit Euler method is used to approximate the differential equations.

In Fig. 1 the results of the simulation are compared with the public recorded data. The vertical axis shows the data as fraction of the total population. The two plots show two different periods, 20 days and 40 days after the first recorded case, for better visibility. In the first 20 days of the disease spread the agreement of the simulated and the real data on the infected individuals is clear, and the trend can be described approximately by an exponential function, or by a polynomial of the 3rd order, given
by Eq. (3) later in the text. In the next 20 days, the real data show more-or-less linear increase, while the simulation still predicts exponential growth. The change in trend of the real data is probably caused by strict measures for disease-spread prevention (contact prevention, social distance, isolation) which are imposed relatively early, about 7–15 days after the first recorded case. Without these measures the development of the number of infected individuals would be much faster and probably it would also follow exponential-growth pattern. The overall disease development predicted by the SIR model for a theoretical case corresponding to the conditions under which this model is derived and the assumptions applying to the adopted model parameters is given in Fig. 2. According to these results, the maximum number of the infected individuals, about 64.5% of the total population, would be reached after 65 days. After that an asymptotic decay is seen. The currently available data (in April 2020, 45 days of disease development) show that the real number of infected individuals still increases following nearly linear pattern.

According to the data given in [6], the total number of infected persons in Bosnia-Herzegovina after 40 days of disease development (on April 14th, 2020) is 1037, which delivers a rate of about 26 newly infected persons per day on average, or about $6.8 \times 10^{-6}$ new infected individuals per day as a fraction of the total population.

According to the same data, the increase of the number of infected individuals in the first 20 days can be described approximately by a polynomial of 3rd degree:

---

**Fig. 1** Number of the infected and the recovered individuals in Bosnia-Herzegovina: comparison of the SIR model with the real recorded data.

**Fig. 2** Overall development of the number of infected and recovered individuals in Bosnia-Herzegovina according to the SIR model.
\[ I \approx 0.03663t^3 - 0.40647t^2 + 1.91293t + 0.12343, \quad (3) \]

where \( t \) is the time in days. If we consider the entire country as the system under observation which is isolated from its environment, so there is no “diffusive” or “convective” transport (surface integrals in Eq. (2) are neglected, since transport across the boundaries is neglected), the source term for this period can be estimated from the approximate equation:

\[
\rho c \frac{\partial \phi}{\partial t} V = c \frac{\partial I}{\partial t} \approx q^+ V \Rightarrow q^+ \approx \frac{3 \cdot 0.03663t^2 - 2 \cdot 0.40647t + 1.91293}{55130} c. \quad (4)
\]

After the first 20 days, the available statistical data reveal quite linear increase of the number of infected persons, which implies that the source is constant, i.e. does not vary with time or with the number of infected persons. The value of the source is estimated to be:

\[
\rho c \frac{\partial \phi}{\partial t} \approx q^+ = 7.9 \times 10^{-4} \frac{1}{\text{km}^2 \text{day}}, \quad (5)
\]

with \( \rho \) equal to 69 persons/km\(^2\) and \( c \) set to 1 for simplicity.

This estimation presumes uniform distribution of the source \( q^+ \) over the entire area of the country. In reality, the source is not uniform, but concentrated in a certain number of locations (typically, large cities and towns). In simulations conducted in this work, the source term \( q^+ \) is defined in the distinct circular regions, defined around 7 cities: Banja Luka, Zenica, Mostar, Konjic, Bihać, Sarajevo, and Tuzla, considered to be important due to the relatively large number of infected persons at the beginning of the disease, as well as due to the relatively high population density. In agreement with previously given conclusions on character of the source term for the entire country, the source term in the affected regions around the mentioned 7 cities is defined as variable in time following the expressions given in Eq. (4) for the first 20 days, and constant after that. The values of numerical coefficients in these expressions are, however, adapted to keep the same value of the total source in the entire country. The magnitude of the sources is not the same for each city. It is distributed in accordance with the public, registered data in the beginning period of the disease by assignment of the appropriate weighting factors.

The destruction term \( q^- \) in Eq. (2) can describe reduction of the number of infected persons, for example including the number of recovered individuals. In this work, however, it is not taken into account since the attention is paid to the progress of infection.

The boundary conditions are given as Dirichlet boundary conditions, described by the number of the infected persons. Here, it is assumed to be equal to zero along all country borders.
Figure 3 shows the computational mesh covering the total area of Bosnia-Herzegovina. The mesh resolution is increased in the regions with higher expected concentration of infected individuals (left). The calculated results obtained solving Eq. (2) without “convective” and “diffusive” transport by the described finite-volume method are given as well, denoted by “numerical method” (right). By neglecting the convective and the diffusive term in Eq. (2), the model simulates the case with completely prevented local contacts among individuals and suppressed distant motion, which would otherwise trigger undirected and directed spatial spread, respectively. Only the increase of the number of infected individuals caused by local sources is obtained, as expected. The simulation results are compared with the results of the previously described SIR model setup as well as with the real data. Their agreement with the real data is obvious. This is expected, since the source term in Eq. (2) is fit to the available real data set, while in SIR model it is defined as proportional to the product of the number of susceptible and infected people in the whole domain. The idea behind this procedure is to calibrate the model to a real case in its beginning stage, in order to be able to provide results of later development as accurate as possible.

The spatial distribution of the number of infected individuals calculated for the cases: (a) without “convection” and “diffusion”, (b) with “diffusion” only, where the diffusion parameter (local transmission capability) is set to 10, and (c) with “diffusion” and “convection” driven by fictitious velocity, randomly generated to fit transportation capacities from 0 to 100 individuals per day, per km of the boundary width are shown in Figs. 4, 5 and 6, respectively. In all cases, the source functions are the same, as previously discussed. The calculated concentrations of the infected individuals \( \phi \) (the number of infected individuals divided by total population) for all three cases are shown by color spectra corresponding to the logarithmic scale, from \( 10^{-6} \) to \( 10^{-1} \), for three instants of time: after 30, 60 and 90 days.

In case (a) all regions with non-zero sources remain isolated, since there are no mixing processes which would be caused by “convection” or “diffusion”. The concentration of the infected individuals increases with time due to the existence of
Fig. 4 Spatial distribution of the concentration of the infected persons, case (a)
Fig. 5  Spatial distribution of the concentration of the infected persons, case (b)
Fig. 6  Spatial distribution of the concentration of the infected persons, case (c)
sources. The total number of infected individuals (the integral of concentration over population in the considered area) and herewith the average concentration over the entire area follow the corresponding curve in Fig. 3 (right).

In case (b) a certain spreading of the regions with sources is seen which increases with time. Some of the regions, which are separated at the beginning but close to each other, merge after sufficiently long time. Due to increase of the area covered by the infected persons, the local maximum concentration in the infected regions is not as high as in case (a).

Due to transportation effects across the local, inner boundaries (in randomly generated directions), the spreading in case (c) is not smooth and not equally distributed in all directions. The covered area seems to be larger than in case (b), and is larger than in case (a). Correspondingly, the local maximum concentration is reduced in the most parts. However, at some places, the additional local maxima of concentration may arise, induced by transported infected individuals, such as in the north-west area, quite close to the border. Depending on the observed level of the concentration of infected individuals, merging of the infected areas may be seen. In the presented example, the convective transport is calculated using randomly generated velocities, and herewith transportation directions and intensity at the beginning of simulation.

5 Conclusions

In this paper, variation of species in time and space within a given population is calculated solving time-dependent convection–diffusion equation using a finite-volume method. The simulation is demonstrated in analysis of spread of individuals infected by SARS-CoV-2 virus, and herewith spread of COVID-19 disease in Bosnia-Herzegovina. Unlike usually used models based on ordinary differential equations only, the mathematical model presented here includes convective and diffusive terms describing directed remote and undirected intermediate transport of infection, respectively. This choice is motivated by capability of these terms to describe variations of the observed variable in space (species concentration i.e. fraction of infected individuals in the given population).

The presented results are plausible. The available public data on disease spread in the beginning period are used to calibrate the source terms. The calculated integral values of the number of infected individuals show a good agreement with the publicly reported values. The results show different responses in spatial spread for three different cases, illustrating scenarios of complete isolation, undirected intermediate contacts, and arbitrary contacts including transportation effects. As expected, the results show clearly larger spread in case of allowed contact/communication between the affected regions, undirected or directed. More detailed and accurate analysis of spatial spread requires exact input data, such as distribution of transportation directions and intensity, or spatial variation in population density, so further research is required in order to collect the necessary data for appropriate modeling of these effects. Additionally, in future research the source and the sink term (production and
destruction terms) could be expanded in order to account for other effects, such as variation of number of recovered individuals, birth rate, deceased individuals etc., or the present model can be combined with one of the SIR-like models in order to take these quantities into account.

References

1. Kermack, W.O., McKendrick, A.G.: A contribution to the mathematical theory of epidemics. Proc. R. Soc. **115**(772), 700–721 (1927)
2. Hethcote, H.W.: The mathematics of infectious diseases. SIAM Rev. **42**(4), 599–653 (2000)
3. Vynnycky, E., White, R.: An Introduction to Infectious Disease Modelling. Oxford University Press, Oxford (2010)
4. Ferziger, J.H., Perić, M.: Computational Methods for Fluid Dynamics. Springer, Berlin (2002)
5. Torlak, M., Hadžiabić, V.: Solving linear wave equation using a finite-volume method in time domain on unstructured computational grids. In: Avdaković, S., (ed.) Advanced Technologies, Systems, and Applications III, Proceedings of the International Symposium on Innovative and Interdisciplinary Applications of Advanced Technologies (IAT 2018), vol. 1, Lecture Notes in Networks and Systems, vol. 59, pp. 347–356, Springer (2019)
6. https://github.com/CSSEGISandData/COVID-19/tree/master/csse_covid_19_data. Access date Apr. 19 2020