Diffusion as a First Model of Spread of Viral Infection

Paulo H. Acioli*

Department of Physics and Astronomy,

Northeastern Illinois University, Chicago, IL 60625

(Dated: March 26, 2020)

Abstract

The appearance of the coronavirus (COVID-19) in late 2019 has dominated the news in the last few months as it developed into a pandemic. In every mathematics and physics classroom, instructors are using the time series of the number of cases to show exponential growth of the infection. In this manuscript we propose a simple diffusion process as the mode of spreading infections. This model is less sophisticated than other models in the literature, but it can capture the exponential growth and it can explain it in terms of mobility (diffusion constant), population density, and probability of transmission. Students can change the parameters and determine the growth rate and predict the total number of cases as a function of time. Students are also given the opportunity to add other factors that are not considered in the simple diffusion model.

I. INTRODUCTION

The end of the year of 2019 and beginning of 2020 has been dominated by the spread of the coronavirus (COVID-19). The disease started in the Hubei province in China in December 2019 and by early January 2020 started to spread. It grew very rapidly, triggering responses from the Chinese and other governments. On March 11, 2020 the World Health organization declared that COVID-19 was then characterized as a pandemic.[1] The situation triggered different responses from different governments. In The United States, many colleges took the initiative to start their own social distance programs, including sending all students back home, extending spring breaks, and ultimately moving all classes to distance learning.[2,3] Many cities and states followed up with shelter at home mandates.[4] The situation in many countries became alarming due to the exponential growth of new cases and deaths.[5,6]

One of the consequences of this pandemic is that instructors at most colleges and universities started to monitor and model the data, using this as a teachable moment for the students and colleagues. The first step in modeling the data is just to plot the number of cases as a function of time and show that it exhibits an exponential growth. For beginners it is important to introduce them to the logarithmic scale, where the plot becomes a straight line and students can extract the exponent and realize that the fit to the US data on March 20, 2020 shows that the number of cases doubles every 2.4 days. Fits to the data for Chicago and New York City show that these cities are having similar growth rates as seen in FIG. 1, however the total number of cases in Chicago is smaller. One can speculate that this might be due to difference in population densities in these cities.

In this work we propose to look at some of the factors affecting the spread of viruses using a simple diffusion model in which each individual in a population is treated as Brownian[9] particle with diffusion constant D. Also added to this model is the incubation period of the virus and a probability of transmission of the virus if individuals are closer than a certain distance. This model is to be used as a project in a computational physics course and verify if it adequately predicts the exponential growth of the number of cases, as well as if the population density, mobility, and probability of transmission play roles on the percentage of the population that will be infected as a function of time. Students, will be asked to write a code, analyze the output for different sets of parameters, and write a critical analysis of its predictive effectiveness.

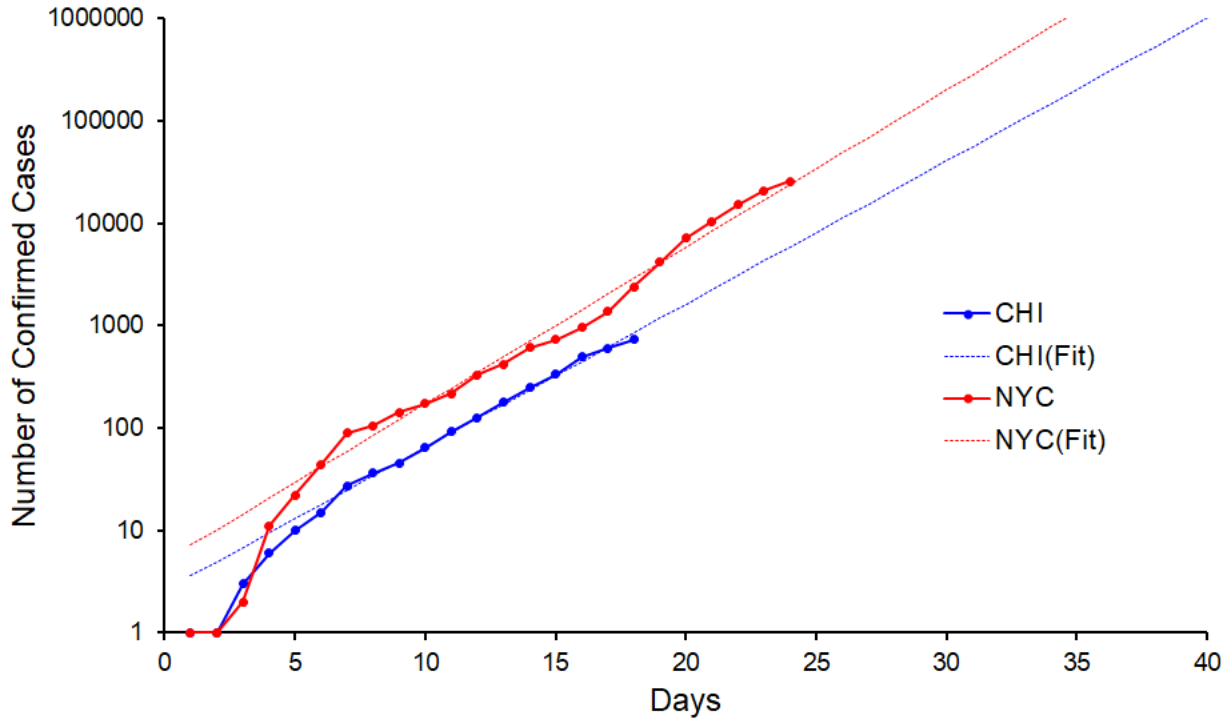

FIG. 1. COVID-19 cases in Chicago (CHI,blue)[7] and New York City[8] (NYC,red). The dashed lines are the best fit to the date after the initial spike in cases. The x-axis is the number of days from the date that the first case in each city was detected.

II. COMPUTATIONAL PROJECT DETAILS

In this section we propose a project to be implemented in a computational physics course to study the spread of infectious diseases as a simple diffusion. The benefits of this project is that its implementation is simple, but it can lead to a qualitative understanding on how diseases are spread and it can also allow the student to understand factors that can affect it. The main ingredients of the model are: individuals are considered particles that obey a Brownian diffusion process; each individual will have three possible states, healthy, sick (contagious), and cured; a healthy individual has a probability of getting infected if its distance to a sick individual is smaller than a certain threshold; the incubation and sickness periods are the same; once an individual gets cured, it cannot be infected or contagious again. VPython (or GlowScript)[10] is the language of choice as it allows for a real time visualization of the infection spread and it will allow for fast simulations of small populations on a laptop.

In the assignment, students will implement the code, analyze the data generated, critique the initial assumptions, and propose improvements for more realistic simulations. Below we describe the algorithm starting with the standard diffusion equation.

The diffusion equation in 2D

$$\frac{\partial f(x,y,t)}{\partial t} = D\left[\frac{\partial^2 f(x,y,t)}{\partial x^2} + \frac{\partial^2 f(x,y,t)}{\partial y^2}\right] \tag{1}$$

is used to study many phenomena[11–17,19? ?] from diffusion inside the nucleus to population dynamics to solving the Shrödinger equation. In this work we will use it for the diffusion of individuals, treated as particles, over closed boundaries subject to contamination of a viral infection. The normalized solution to Eq. (1) for a single particle is given by

$$f(x,y,t) = f_0 \frac{1}{\sqrt{4\pi D t}} e^{-\frac{x^2+y^2}{4Dt}}. \tag{2}$$

Therefore one can simulate the diffusion of a particle from its previous position by generating a Gaussian distribution of zero mean and variance $\sqrt{2Dt}$. For a system on N non-interacting particles with the same diffusion constant we use Eq. (**??**) for each particle at each simulation time step.

The next ingredients in the simulation will be the population density ρ, the number of habitants (particles in the simulation cell) $Npop$, the diffusion constant D, the number of simulation steps $Nstep$, the time step dt, the incubation period (t_{inc}), the transmission radius, and the probability of transmission from an infected to a healthy individual ($prob$). All of these variables are set at the beginning of the simulation. We preset that the total simulation time consists of 90 days and that each time step is 0.01 days, therefore each simulation takes 9000 steps. The algorithm is described below:

1. Input $Npop$, $Nstep$, D, dt, t_{inc}, $prob$, r_{transm}
2. Calculate the size of the square cell as $L = \sqrt{(Npop/\rho)}$)
3. Initialize the initial population
4. Choose a fraction of the initial population to be infected, and set timer for the sickness (t_{sick})
5. Loop over $Nstep$

6. Move all individuals according to the Gaussian distribution (Eq. (**??**)).

7. Compute the distance between each healthy and infected individuals

   If the distance is less than r_{transm}, the healthy individual becomes sick with probability $prob$

8. Subtract the sickness timer by dt

9. If $t_{sick} < 0$ the sick individual gets cured

In order to generate more accurate statistics we suggest the students run simulations with the same initial parameters multiple times, between 20 and 100, depending on the size of the system and the speed of the student's computer. We suggest that the smallest population to contain 100 individuals, as even on sick individual corresponds to an initial infected population of 1%. Depending on how much time the students have to complete the analysis of the project, they can use a population of 1000, being aware that each individual 9000 steps simulation can take up to a 30minutes. They can speed up the process by using larger time steps, however, they must test if the results of an individual simulation with the same diffusion constant and different time steps lead to similar outcomes.

III. AN EXAMPLE OF A PROJECT

In this section we provide an example of a project that students could pursue. In this project, student will study the differences of the infection proliferation in New Your City (NYC) and in Chicago. New York City was chosen because its is experiencing a very rapid growth on the number of cases and preventive measures such as shelter at home were taken by the city and state governments. Chicago was chosen as a local connection to Northeastern Illinois University (NEIU) students, and is also experiencing an exponential growth in the number of cases, however, is still in the earlier stages and has also been affected by shelter at home mandates. We must be mindful that current model will not be able to study the effects of the measures that the government are taking, but it might be able to justify their need.

The population density of New York city is 10,194 people/km^2,[21] while Chicago's is 4,665 people/km^2.[22] It is natural to offer the students the hypothesis that if all other variables

are the same that the spread of infection in New York City will be faster than in Chicago and that the number of cases will be much larger for the same period of time after the first case. With this hypothesis alone, students should be able to generate a enough data for the project. They can also discuss if the mobility (diffusion constant) should be the same in both cases and study the effect of mobility in the spread of the disease. In the proposed model the period of incubation is the same as the period of sickness, students can be offered the option to modify this assumption. The model also assumes that once cured, an individual cannot get sick again, and will not be able to spread the disease. They can discuss modifications to the model to incorporate relapse. It is clear that with this very simple model and this very limited two-city project, they can perform a very thorough study that can give us some qualitative understanding of infectious disease spread.

We start with preliminary simulations of populations of 100 individuals and 1 sick individual chosen at random. The initial diffusion constant was chose at $D = 100\text{m}^2/\text{day}$ with a time step $dt = 0.01\text{day}$. Therefore the variance of the gaussian distribution is 1. In this case the number of simulation steps is 9000 for each individual simulation is 27s in a laptop with Intel(R) Core(TM) i5-8250U CPU @ 1.60GHz. Therefore 50 simulations will take about 22.5 minutes of computation. We assume that the radius of contamination is 2m, and that the probability of contamination is 10% for each time step (14.4 min) that a healthy individual is within the radius of contamination of a sick individual. The incubation period is taken as 14 days. Below are some of the results of these simulations.

In FIG. 2 we present the results of the simulation for New York City with the set of parameters in the previous paragraph. As one can see, under this assumption 64% of the population is sick on day 21 and 99% of the population will be infected after 50 days. One can see that after day 26 the number of sick people starts to decline. However, if the death rate is similar to what has been observed on the COVID-19 pandemic, about 6% of the population of NYC would perish, and the numbers would be even worse, since no major city in the world would be able to have hospital beds for 64% of its population. In the figure we also observe that the rate of cure is follows the number of infected individuals with a lag time of 14 days, which is the incubation/sickness period.

In FIG. 3 we reduce the diffusion constant to $D = 50\text{m}^2/\text{day}$ and keep all other variables the same. One can see that under these conditions, and starting with 1% of the population already infected,that after 76 days 77% of the population will be infected. However, the

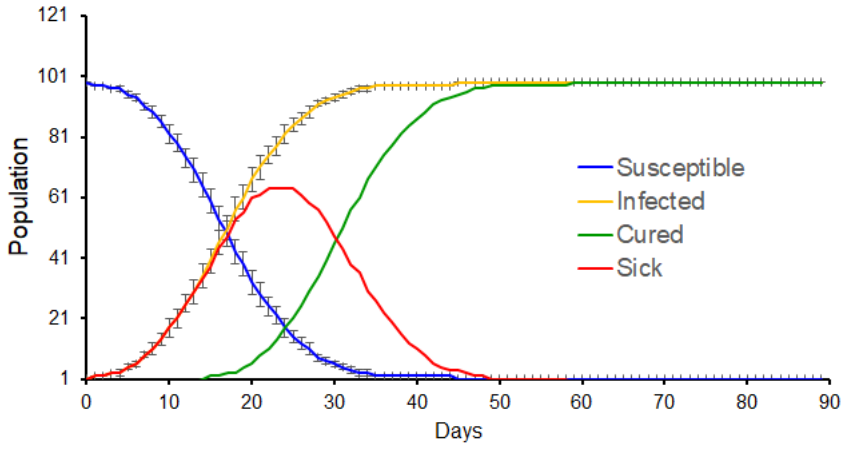

FIG. 2. Average of 50 90-day simulations of the spread of a virus in a population of 100 individuals in a square cell with the population density of NYC ($\rho = 0.012$ people/m^2). $D = 100$m^2/day,$prob = 0.1$, $dt = 0.01$day.

peak of the sick population is at 20%. Although, still catastrophic this scenario is much more favorable than the one presented in FIG. 2. That means that reducing the mobility of the population even further can lead to more manageable outcomes. Since, this manuscript is not meant to be a study of the spread of the virus, we will let to interested instructors to assign to their students to do a more thorough study of the effect of mobility in the spread of the disease. In the next two cases we will repeat the previous two simulations will all the same parameters, but the population density, that will be changed to $\rho = 0.0047$ people/m^2, representative of the city of Chicago.

The results of the simulation for a population of 100 individuals in a region with the population density of Chicago are shown in FIG. 4. For this set of parameters 28% of the population will be infected after 83 days. However, the peak of the number of sick people is at 8% and it happens about 26 days after the first 1% of the population is infected. This number is still far above of what can be handled by the current number of hospital beds and emergency equipment available. However, it does support our initial hypothesis that under the same conditions, one would expect a slower growth rate in a less dense population.

To confirm that mobility is still a big factor we repeat the simulation for the sample representative of Chicago with $D = 50$m^2/day. The results are shown in FIG. 5. Under these conditions, only 7% of the population will be infected with the peak of sick people

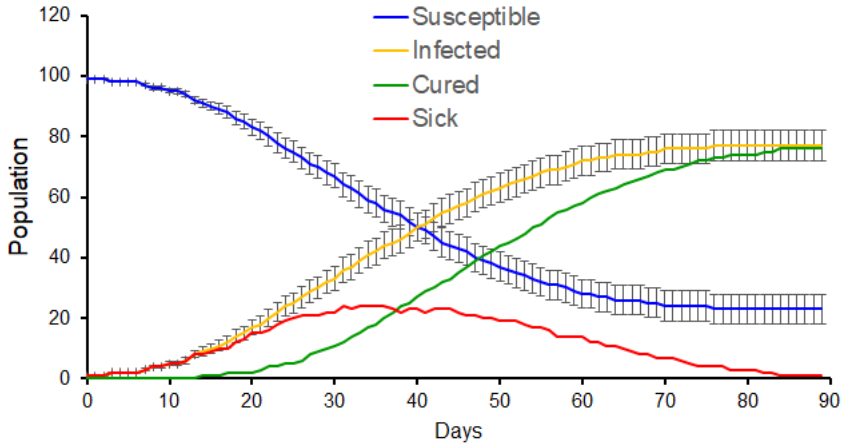

FIG. 3. Average of 50 90-day simulations of the spread of a virus in a population of 100 individuals in a square cell with the population density of NYC ($\rho = 0.012$ people/m^2). $D = 50\text{m}^2/\text{day}$,$prob = 0.1$, $dt = 0.01\text{day}$.

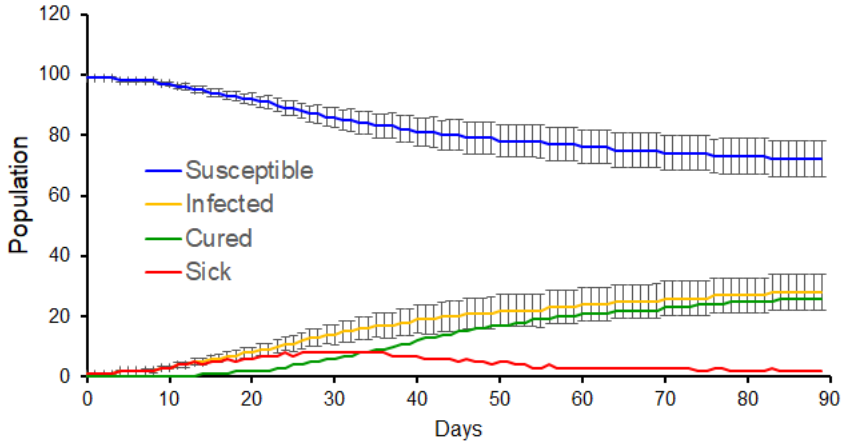

FIG. 4. Average of 50 90-day simulations of the spread of a virus in a population of 100 individuals in a square cell with the population density of CHI ($\rho = 0.0047$ people/m^2). $D = 100\text{m}^2/\text{day}$,$prob = 0.1$, $dt = 0.01\text{day}$.

at 3% of the population or roughly 65,000 people sick at one time. This shows that under this model large mobility in high density population wold lead to more severe spread of the infection. A more black box option would be to resort to the NDlib Python library that models diffusion processes over complex networks.[23]

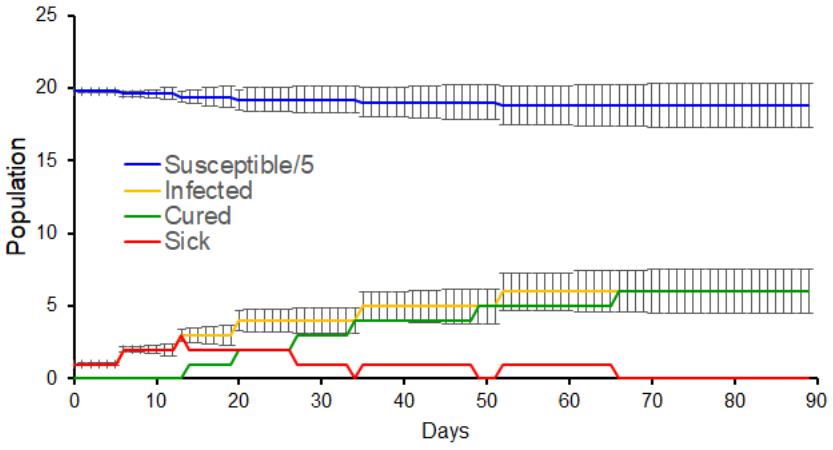

FIG. 5. Average of 50 90-day simulations of the spread of a virus in a population of 100 individuals in a square cell with the population density of CHI ($\rho = 0.0047$ people/m^2). $D = 50$m^2/day,$prob = 0.1$, $dt = 0.01$day.

This project can help students understand particle diffusion and extrapolate some of its findings into a timely phenomenon such as viral infection. It is interesting to note that these parameters do not tell the whole story. For instance, FIG. 1 is showing that the growth rate of case in the states of New York and Illinois are roughly the same. However, the model for the largest cities in these states predict very different growth rates. Here is an opportunity for students to discuss the shortcomings of the model and/or the initial assumptions. For instance, one can think that the mobility in a denser population is smaller than that of a less dense one. The combination of these two factors could lead, within this simple model to identical results. Looking at FIG. 5, student might ask themselves will the graphs are not smooth. The answer of course, lies in the fact that we are simulating a small population and the changes by one individual seem to be discrete. One way to solve this issue is to increase the population. On FIG. 6 we plot the results of the simulations for a population of 1000 with a population density of ($\rho = 0.0047$ people/m^2) and an initial population of sick people representative of 10 (1% of the total population). One can see that the data is smoother and although the trends are similar to those shown in FIG. 5, at the end of 90 days 20% of the population is infected, while in the 100 population case this number if 6%. The peak of people that are sick peaks at roughly 5% in the 1000 population at the 40th day, while the peak for the 100 population is at 3% at the 14th day. This is an important example on the

need to scale the simulations to make sure the predictions are converged. Students should be directed to vary the populations size until the predictions agree. For the sake of keeping the computational time manageable for a short term project we suggest 100, 200, 400, 800, 1600.

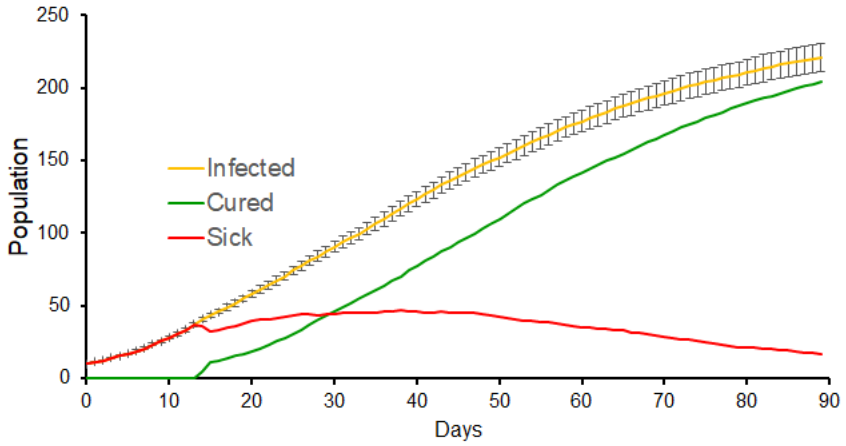

FIG. 6. Average of 50 90-day simulations of the spread of a virus in a population of 1000 individuals in a square cell with the population density of CHI ($\rho = 0.0047$ people/m^2). $D = 50\text{m}^2/\text{day}$,$prob = 0.1$, $dt = 0.01\text{day}$. We start with 10 sick individuals on day 1.

To finalize, we show snapshots of the simulation cell and its population in FIG. 7. White spheres represent the susceptible people, red represent sick, and green are those individuals that recovered. This window is very useful for a quick analysis of what is happening and can show how the infection is spread over time. We found it useful to change the radius of the individuals so that they are visible as the population increases, and the cell size on the screen remains the same.

IV. CONCLUSION

This project on itself seems simple enough that it can be implemented and analyzed in a first course on computational physics. However, it is rich enough that can lead to deep analysis of a spread of viral infection. In addition, it can help students understand the phenomenon of particle diffusion. In the example presented the simulation was limited to the same 10% change of infection, the same time step, the size of the population was fixed,

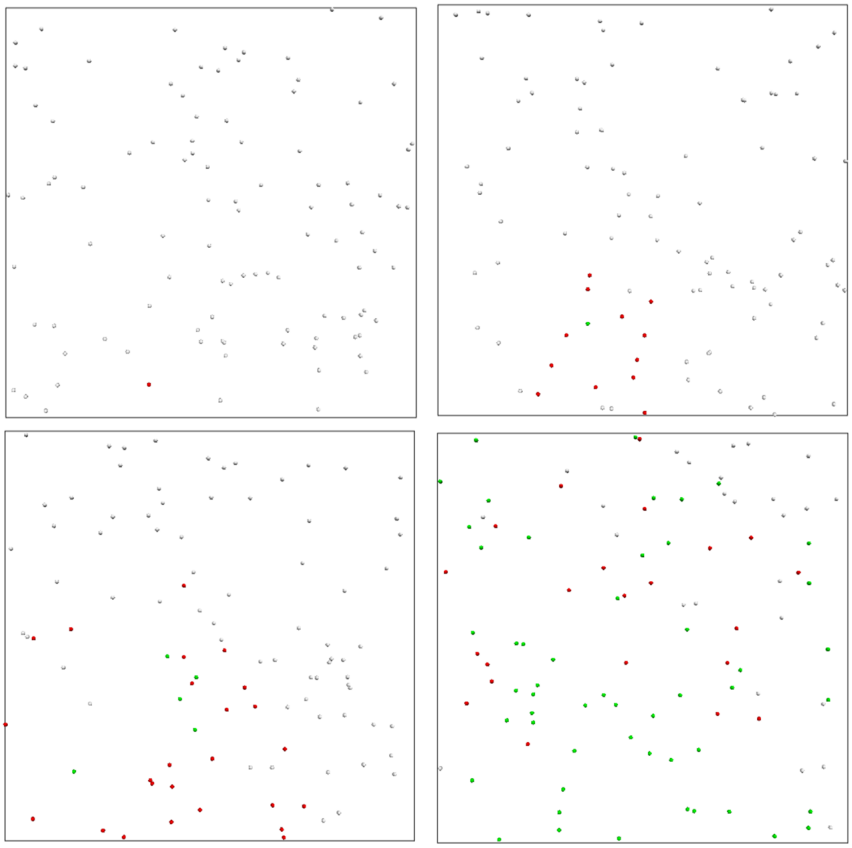

FIG. 7. Snapshots of one of the simulation 90-day simulations of the spread of a virus in a population of 100 individuals in a square cell with the population density of NYC ($\rho = 0.012$ people/m^2). $D = 50\text{m}^2/\text{day}$,$prob = 0.1$, $dt = 0.01$day. a) Initial Distribution. b) 14 Days. c) 28 Days. d) 50 days.

and the length of the disease was the same as the incubation period. Under these conditions students will have a larger number of parameters to change and verify if they can reproduce data readily available in the news. In addition, they will have the opportunity to determine whether or not the growth in the number of cases follows an exponential trend and notice that once it reaches a certain threshold if it will follow the expected logistic behavior.

Once they explore the model as its presented they can suggest modifications. For instance, there could be an incubation period, and then a symptomatic period, where they will be contagious, and then be cured or if they want to be realistic add a death rate. They could also add a drift term that will attract the population to concentrate in certain areas. After they explore all these modifications to the particle diffusion model they can move on to more sophisticated simulations such as the one in ref.[24], where they incorporate a probability of

infection that depends on the mobility between regions and instead of individuals they have population densities assigned to different cells on a grid.

In conclusion, we hope to have convinced the reader that a simple diffusion model can be used to qualitative explain the spread of disease and even in same case quantify it. It is our judgment that it will allow students to work on a problem that are directly affecting them and that these simulations will help them offer valuable insight on the issue.

ACKNOWLEDGMENTS

I would like to thank Dr. Greg Anderson for sparking the discussion of using the COVID-19 pandemic as a teachable moment and Dr. Orin Harris for inspiring the idea of comparing the data for different populations.

---

* p-acioli@neiu.edu

[1] World Health Organization Director-General's March 11, 2020 COVID-19 Media Briefing. (https://www.who.int/dg/speeches/detail/who-director-general-s-opening-remarks-at-the-media-briefing-on-covid-19—11-march-2020)

[2] https://www.wsj.com/articles/coronavirus-prompts-colleges-to-send-students-home-11583862936.

[3] https://www.usatoday.com/videos/news/2020/03/14/coronavirus-colleges-spring-break/5046501002/

[4] https://www.usatoday.com/story/news/nation/2020/03/21/coronavirus-lockdown-orders-shelter-place-stay-home-state-list/2891193001/

[5] https://coronavirus.jhu.edu/map.html

[6] https://www.who.int/emergencies/diseases/novel-coronavirus-2019/situation-reports

[7] https://www.chicago.gov/city/en/sites/covid-19/home.html

[8] https://www.ny1.com/nyc/all-boroughs/news/health-and-medicine/coronavirus-blog

[9] Feynman, R. (1964). "The Brownian Movement". The Feynman Lectures of Physics, Volume I. pp. 411.

[10] https://vpython.org/

[11] F. A. Williams, Amer. J. Phys. **26**, 467 (1958).

[12] D. C. Kelly, Amer. J. Phys. **36**, 585 (1968).

[13] Karmeshu, and L. S. Kothari, Amer. J. Phys. **40**, 1264 (1972).

[14] C. Domb and E. L. Offenbacher, Amer. J. Phys. **46**, 49 (1978).

[15] J. DArruda, and E. W. Larsen, Amer. J. Phys. **46**, 392 (1978).

[16] A. M. Albano, N. B. Abraham, D. E. Chyba, and M. Martelli, Amer. J. Phys. **52**, 161 (1978).

[17] S. Redner and P. L. Krapivsky, Amer. J. Phys. **67**, 392 (1999).

[18] K. Ghosh, K. A. Dill, M. M. Inamdar, E. Seitaridou, and R. Phillips, Amer. J. Phys. **74**, 123 (2006).

[19] D. R. Spiegel and S. Tuli, Amer. J. Phys. **79**, 747 (2011).

[20] T. Pang, Amer. J. Phys. **82**, 980 (2014).

[21] "NYC Population: Current and Projected Populations". www1.nyc.gov. Retrieved May 2, 2019.

[22] https://www.opendatanetwork.com/entity/1600000US1714000/Chicago_IL/geographic.population.density?yea

[23] G. Rossetti, L. Milli, S. Rinzivillo, A. Sîrbu, D. Pedreschi, and F. Giannotti Int. J. Data Sci. and Anal. **5**, 61(2018).

[24] https://towardsdatascience.com/modelling-the-coronavirus-epidemic-spreading-in-a-city-with-python-babd14d82fa2