MODELS IN DEMOGRAPHY

Christopher Ableitinger

Center for Teacher Education , University of Vienna

Abstract

In this article, we present a way to use GeoGebra for working on so called Leslie models. Biomathematicians use these models to describe the demographic development of populations. We have used original statistics on the population of the city of Vienna to provide realistic data. By following the presented tasks students gain insights into the field of demography by starting with the familiar exponential growth model that is successively advanced to a simple Leslie model that takes into consideration the age structure of the population. GeoGebra serves as an ideal tool for implementing these models and allows for some predictions for the future development of the population being considered.

Keywords: Biomathematics, GeoGebra, Spreadsheets, Iterative processes, Growth models

1 INTRODUCTION

In classrooms the work on demographic models is usually limited to the exponential model. In some cases the logistic model is taken into account as well. In this article we want to show how the use of GeoGebra can broaden the view on this interesting and relevant topic. We will present concrete tasks for students at level 7–10 that will lead them to a simple version of Leslie models (Ableitinger, 2013; Britton, 2012). The data refers to the population of the Austrian capital Vienna in the past 200 years. If available, you can of course use data of your hometown as well.

The Viennese population increased almost exponentially during the period from 1810 to 1910. This gives you ideal access to iterative processes (Weigand, 2004) that can be simulated with the spreadsheet functions of GeoGebra. Although organizing and managing the use of spreadsheets in class is a rather complex task for teachers (Haspekian, 2005) it provides possibilities to model iterative processes with very few technical and syntactical requirements. The students should be familiar with entering data and formulas in the spreadsheet, with the tool "*Slider*" and basic functions of the graphics window.

2 EXPONENTIAL GROWTH

We presume that the students already know about the recursion formula of the exponential growth model $N_{t+1} = N_t \cdot (1 + a)$, where N_t is the number of people living in Vienna at time t and a is the difference between the birth rate and the death rate of the Viennese population (in the following a is assumed to be a constant).

Task 1

Enter the following data (Table 1) into two columns of the spreadsheet of GeoGebra and visualize it in the graphics window.

Year	1810	1820	1830	1840	1850	1860	1870	1880	1890	1900	1910
Population	225	300	400	470	550	700	900	1165	1430	1770	2085

Table 1. Population of Vienna in thousands (1810-1910). Source: Statistik Austria (2017).

Discussion of Task 1: The Viennese population shows an almost exponential growth, as you can see in Figure 1. For plotting the data entered in the spreadsheet, select the data and use the GeoGebra tool "List of Points". Click on the button "Create".

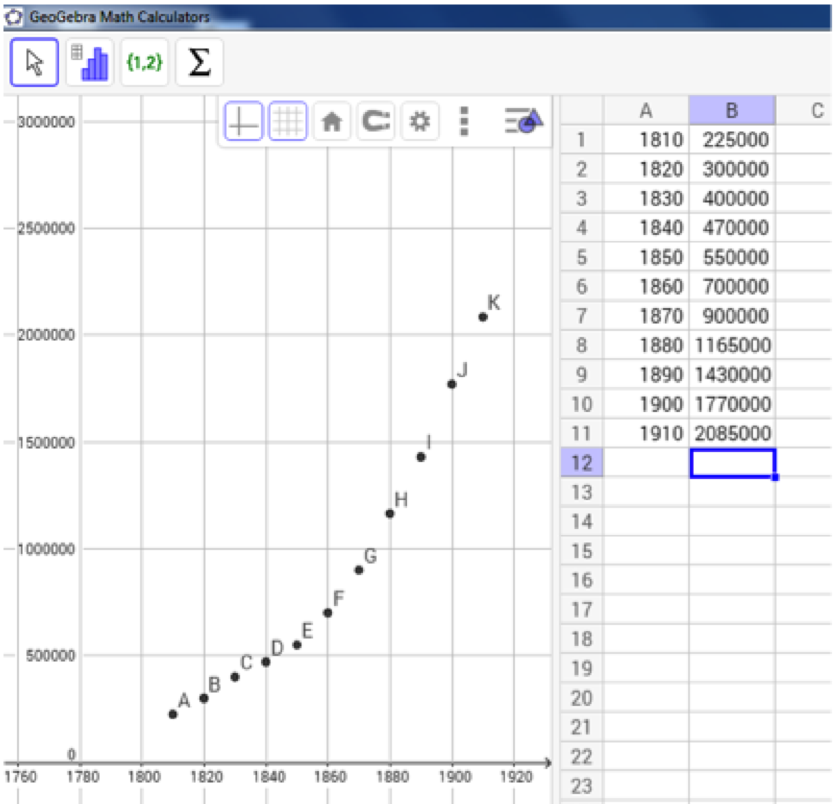


Figure 1. Population of Vienna (1810-1910)

Task 2

Try to approximate the data in Table 1 using the discrete exponential growth model with the recursion formula $N_{t+1} = N_t \cdot (1 + a)$. Use 10 years as a time step. Find the ideal value for parameter a.

Discussion of Task 2: Use the tool "Slider" for varying the value of parameter a. We will start with a population of 225,000 in the year 1810 (cell D2 in Figure 2) and enter `=D2*(1+a)` into cell D3. We then transfer this formula to the cells D4 to D12. As you will notice by trying, the model fits the data best for $a = 0.25$.

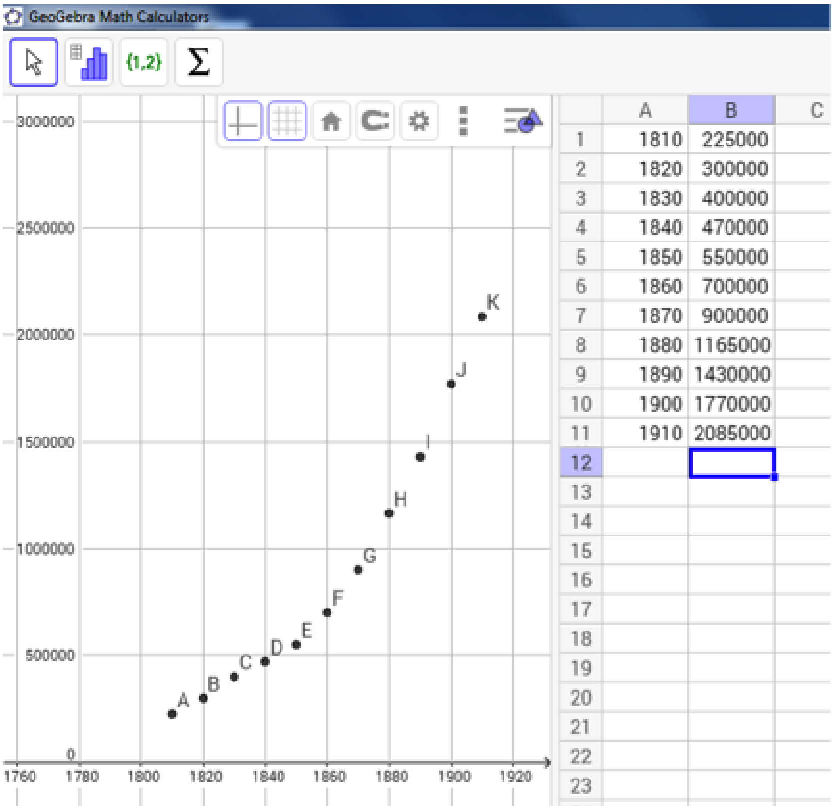


Figure 2. Exponential growth model for $a = 0.25$

Task 3

Imagine if the Viennese population had grown exponentially in the period from 1910 to 2010 as well. How many people would have lived in Vienna in the year 2010?

Discussion of Task 3: Following the assumption of exponential growth ($a = 0.25$) there would have been more than 19 million people in Vienna. But in reality the development of the Viennese population changed considerably after 1910 (Table 2). This is what Task 4 is about.

Year	1910	1920	1930	1940	1950	1960	1970	1980	1990	2000	2010
Population	2085	2100	1920	1770	1600	1625	1620	1530	1500	1550	1700

Table 2. Population of Vienna in thousands (1910-2010). Source: Statistik Austria (2017).

Task 4

Compare the data in Table 2 to the calculations from Task 3. What are the reasons for the deviation from the prediction of the exponential growth model?

Discussion of Task 4: The two World Wars led to a stagnation or rather a decline of the Viennese population in the first half of the 20th century. Afterwards social and political changes led to a decreased birth rate in the Austrian population. Therefore the number of inhabitants did not exceed the threshold of 2 million people again in the considered period of time.

As you have seen the exponential growth model has its limits. There are a lot of aspects that it does not take into account. One of these aspects is the birth rate's dependence on the number of fertile females in the population. We are going to take into account this dependence in the next section.

3 FERTILITY

Task 5

In the exponential growth model we presumed that the increment of the population ($a \cdot N_t$) is proportional to the current population number (N_t). Explain why this is an unrealistic assumption. Find important factors that influence the increase/decrease of the population number.

Discussion of Task 5: First of all constant birth and death rates are unrealistic. Political, social and medical factors affect these parameters. Additionally, migration plays an important role. Finally, the age structure of the population influences the number of births and deaths in the population being considered. We are going to concentrate on this aspect in the following.

Task 6

We want to divide the female population into age groups depending on the fertility of the women in them. Take into account that we want to develop a model for the population of Vienna, i.e. for a city in a developed country. Find a reasonable division and try to justify the restriction to the female part of the population.

Discussion of Task 6: Of course there are a lot of possibilities to divide the female part of the population into age groups. For our purpose it is advisable to use age groups of equal width, for example 20 years. This leads to a model with five age groups:

1. 0 to 19 years: fertility relatively small
2. 20 to 39 years: fertility high
3. 40 to 59 years: fertility relatively small
4. 60 to 79 years: fertility null
5. 80 to 99 years: fertility null

The restriction to the females in the population makes it easier to work with the model afterwards. One does not have to think about the fact that even older men can become fathers. Due to the fact that about half of the newborns in a population are female and the age-related death rates of females and males are comparable to the male part of the population is going to develop quite similarly to the female part. Therefore we can neglect the consideration of the male part of the population in our model.

Task 7

Consider the information about the Viennese population in Table 3.

a) We denote the number of females from 0 to 19 years at time t by A_t, from 20 to 39 years by B_t and so on. We determine 20 years as one time step so that each woman belongs to each age group for only one time step (an initially 15 year old girl is 35 years old one time step later for instance). We choose $t = 0$ for the year 2016. Identify $A_0, B_0, \dots, E_0$ in the table.

b) On average, how many girls does a woman from age group A give birth to in one time step (20 years)? We name this number "fertility rate (of daughters)" and denote it by f_A.

c) We denote the fertility rates for the age groups $B, \dots, E$ by $f_B, \dots, f_E$ accordingly. Calculate $f_B, \dots, f_E$ by using the data(Statistik Austria, 2017) in Table 3.

Age group (age in years)	Number of females in this age group (year 2016, rounded)	Number of newborn girls with mothers in this age group (year 2016, rounded)
A: 0 to 19	172,000	240
B: 20 to 39	288,000	9640
C: 40 to 59	262,000	530
D: 60 to 79	179,000	0
E: 80 to 99	51,000	0

Table 3. Age groups, number of females, number of newborn girls in Vienna (2016).

Discussion of Task 7: The calculation of the fertility rates $f_A, \dots, f_E$ is challenging for students because there are a few aspects they have to consider. We start with the 240 girls that were born by women of age group A in 2016. If we assume this number to be constant over the years we can expect about $20 \cdot 240 = 4800$ newborn girls in one time step with mothers from age group A. This number of 4800 now has to be divided by the number of women in age group A, i.e. $f_A \approx \frac{4800}{172000} \approx 0.028$. The fertility rates of the other age groups are computed analogously: $f_B \approx 0.669, f_C \approx 0.040, f_D = 0, f_E = 0$.

4 LESLIE MODEL

Task 8

Before we continue developing an age-structured model for the Viennese population we make two assumptions to keep the model simple: the fertility rates $f_A, \dots, f_E$ shall be constant over the years and there shall be no mortality in age groups A to D, i.e. all women belonging to age group A in 2016 shall have moved to age group B one time step later (i.e. in the year 2036).

a) Reflect on these two assumptions and explain why their suitability is limited.

b) Compute the number of females in age groups A, B, C, D and E in the year 2036 following the two assumptions.

Discussion of Task 8: It is clear that the fertility rates will change over the years. In the past social changes such as the invention of the birth control pill, emancipation or capitalism has led to a decline of fertility rates. Politics also influence birth rates by setting financial incentives. The second assumption is unrealistic since mortality unfortunately does play a role in all age groups.

It is easy to determine $B_1, \ldots, E_1$ (i.e., the number of women in these age groups in the year 2036) because they arise from A_0, B_0, C_0, and D_0 directly: $B_1 = A_0 = 172000, C_1 = B_0 = 288000, D_1 = C_0 = 262000$ and $E_1 = D_0 = 179000$. The calculation of A_1 is a little bit more complicated. This group is composed of the newborns with mothers in age groups A, B and C. We therefore get $A_1 = A_0 \cdot f_A + B_0 \cdot f_B + C_0 \cdot f_C \approx 208000$.

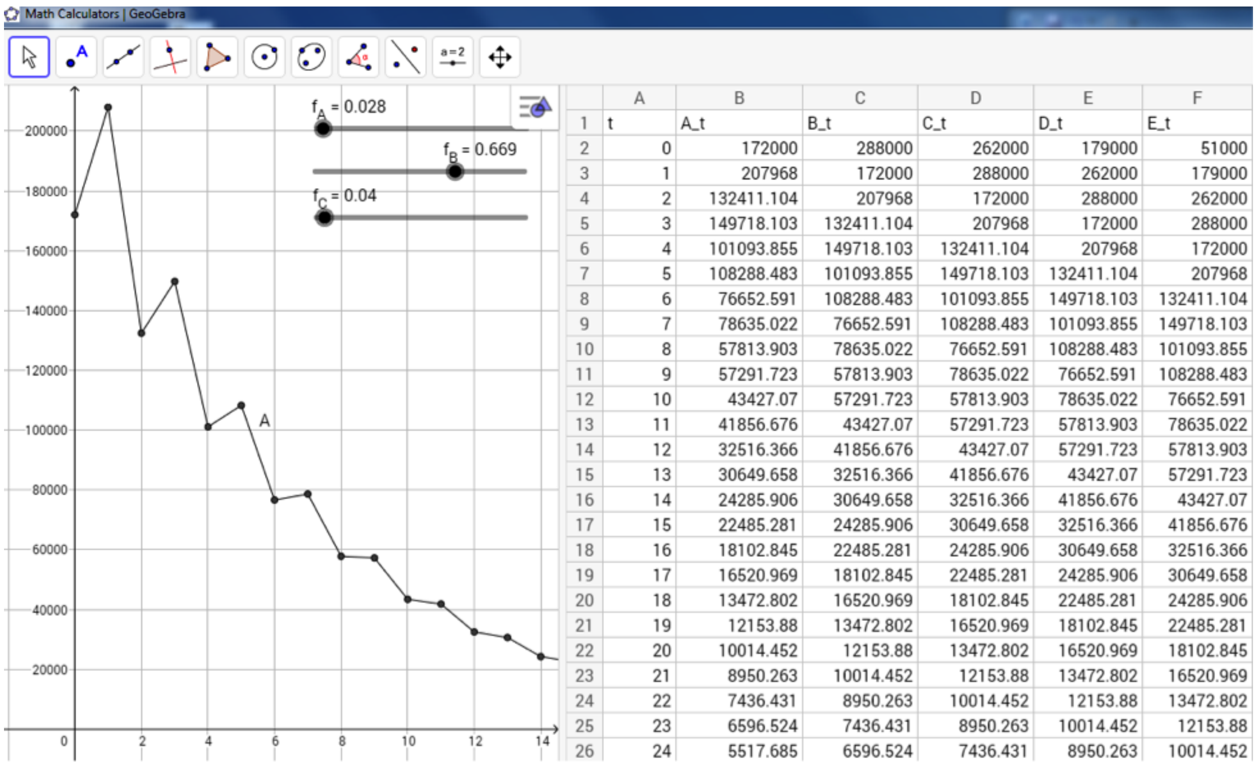
Task 9

Find general recursion formulas for this population, i.e., formulas for computing $A_{t+1}, B_{t+1}, C_{t+1}, D_{t+1}$, and E_{t+1} using A_t, B_t, C_t, D_t, and E_t. Implement these formulas in the spreadsheet of GeoGebra using the values for f_A, f_B, and f_C computed above.

Discussion of Task 9: Following the concrete calculations in Task 8 you get the following system of coupled difference equations:

$$\begin{aligned} A_{t+1} &= A_t \cdot f_A + B_t \cdot f_B + C_t \cdot f_C \\ B_{t+1} &= A_t \\ C_{t+1} &= B_t \\ D_{t+1} &= C_t \\ E_{t+1} &= D_t \end{aligned}$$

After plotting the data for one age group use the command “Polyline” to connect the points in the graphics window. In Figure 3 you can see the simulation of age group A using GeoGebra.



	A	B	C	D	E	F
1	t	A_t	B_t	C_t	D_t	E_t
2	0	172000	288000	262000	179000	51000
3	1	207968	172000	288000	262000	179000
4	2	132411.104	207968	172000	288000	262000
5	3	149718.103	132411.104	207968	172000	288000
6	4	101093.855	149718.103	132411.104	207968	172000
7	5	108288.483	101093.855	149718.103	132411.104	207968
8	6	76652.591	108288.483	101093.855	149718.103	132411.104
9	7	78635.022	76652.591	108288.483	101093.855	149718.103
10	8	57813.903	78635.022	76652.591	108288.483	101093.855
11	9	57291.723	57813.903	78635.022	76652.591	108288.483
12	10	43427.07	57291.723	57813.903	78635.022	76652.591
13	11	41856.676	43427.07	57291.723	57813.903	78635.022
14	12	32516.366	41856.676	43427.07	57291.723	57813.903
15	13	30649.658	32516.366	41856.676	43427.07	57291.723
16	14	24285.906	30649.658	32516.366	41856.676	43427.07
17	15	22485.281	24285.906	30649.658	32516.366	41856.676
18	16	18102.845	22485.281	24285.906	30649.658	32516.366
19	17	16520.969	18102.845	22485.281	24285.906	30649.658
20	18	13472.802	16520.969	18102.845	22485.281	24285.906
21	19	12153.88	13472.802	16520.969	18102.845	22485.281
22	20	10014.452	12153.88	13472.802	16520.969	18102.845
23	21	8950.263	10014.452	12153.88	13472.802	16520.969
24	22	7436.431	8950.263	10014.452	12153.88	13472.802
25	23	6596.524	7436.431	8950.263	10014.452	12153.88
26	24	5517.685	6596.524	7436.431	8950.263	10014.452

Figure 3. Simulation of the model using GeoGebra, time diagram for age group A

Task 10

Interpret the results of Task 9 with regard to content. What does the model predict for the Viennese population? What can one do to prevent a development like this? In what way do the time diagrams for the age groups B, C, D and E look alike?

Discussion of Task 10: Following the predictions of the model the population would reduce in size successively and die out sooner or later. The low fertility rates of the population are the reason for this development. This is a common problem for developed countries. They try to meet this challenge by setting financial incentives for young families or by improving childcare. In some of the countries immigration leads to a stabilization of the population size. The time diagrams for the other age groups look similar to the one for age group A, they are only offset to the right.

5 MORTALITY

The alarming result of Task 9 gets even worse when taking into consideration the mortality in age groups A, B, C and D.

Task 11

Consider Table 4 where you can see the probabilities for a female born in 2016 in Vienna to reach a certain age. Why can't we use the data the way it is presented in Table 4? What kind of data do we need instead and how can we derive it using the data in Table 4?

Age	0	10	20	30	40	50	60	70	80	90
Probability reach age	100	99.6	99.5	99.3	98.9	97.9	95.2	88.5	73.4	35.2

Table 4. Probabilities to reach a certain age in percent (for a female born in 2016 in Vienna). Source: Statistik Austria 2017

Discussion of Task 11: What we can't find in Table 4 is the probability for a woman in a certain age group being alive and therefore reaching the following age group 20 years later. We can't compute these probabilities directly from the probabilities in Table 4. One problem is that a 59 year old woman has a much higher probability of reaching the next age group than a 41 year old woman (who hasn't managed to become 59 years old yet). Therefore we choose a representative person for each age group, i.e., a ten year old girl for age group A and ask for the probability that she will reach an age of 30. The percentage of newborn females to become at least 10 years old is 99.63%, the one to become at least 30 years old is 99.29% (see Table 4). The probability of survival s_A (i.e., the probability for changing from A to B alive) has to meet the following equation: $0.9963 s_A = 0.9929$. This yields a very high probability of survival of $s_A \approx 0.997$. Analogously we can calculate the probabilities of survival for the other age groups: $s_B \approx 0.986, s_C \approx 0.904$ and $s_D \approx 0.398$.

Task 11 is of course one of the most challenging ones in the whole unit. The students may need some hints and support from the teacher. The presented way of deriving probabilities of survival is not the only one. Maybe the students will have other creative ideas.

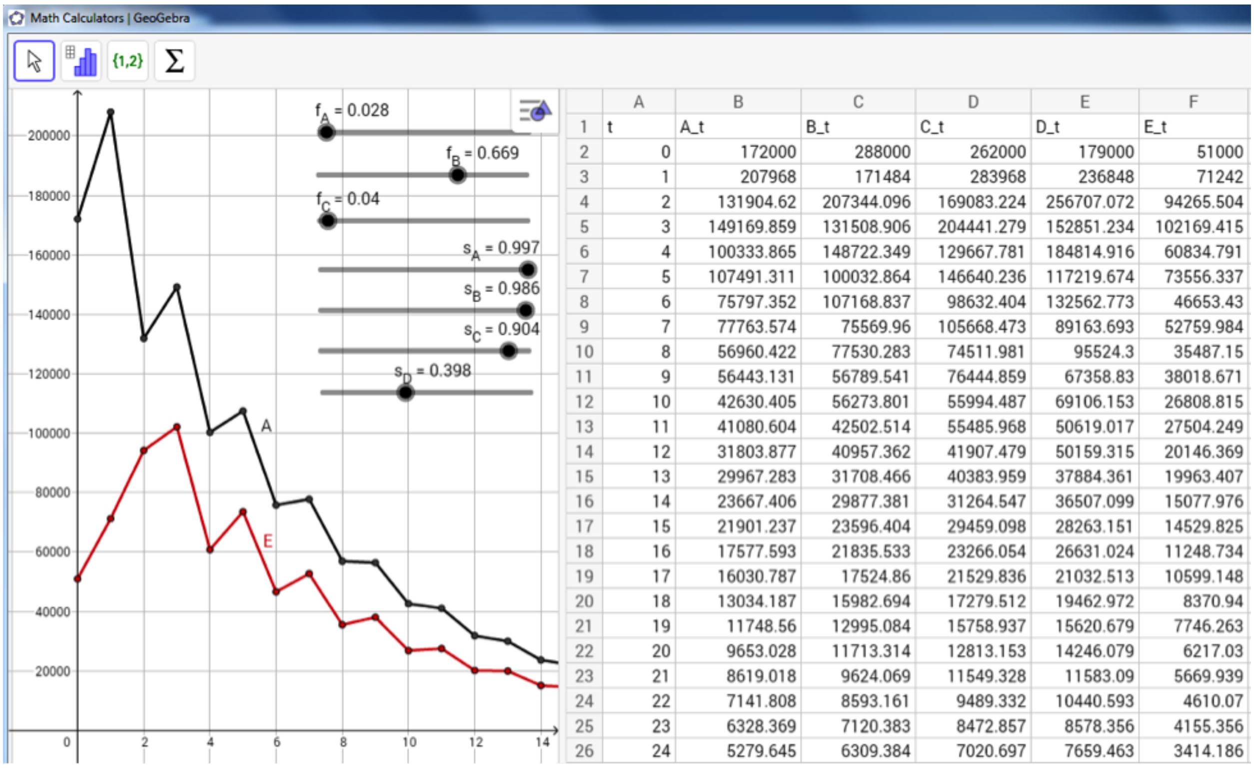
Task 12

Adjust the model from Task 9 by taking mortality into consideration. Implement the new model in GeoGebra and have a look at the time diagrams again. What has changed?

Discussion of Task 12: The system of coupled difference equations looks slightly different from the one in Task 9:

$$
\begin{aligned}
A_{t+1} &= A_t \cdot f_A + B_t \cdot f_B + C_t \cdot f_C \\
B_{t+1} &= A_t \cdot s_A \\
C_{t+1} &= B_t \cdot s_B \\
D_{t+1} &= C_t \cdot s_C \\
E_{t+1} &= D_t \cdot s_D
\end{aligned}
$$

Figure 4 shows the simulation in GeoGebra and the time diagrams for age groups A and E. The population size of age group A hasn't changed compared to the simulation in Task 9, the number of women in age group E on the other hand has decreased drastically because of the mortality involved.



	A	B	C	D	E	F
1	t	A_t	B_t	C_t	D_t	E_t
2	0	172000	288000	262000	179000	51000
3	1	207968	171484	283968	236848	71242
4	2	131904.62	207344.096	169083.224	256707.072	94265.504
5	3	149169.859	131508.906	204441.279	152851.234	102169.415
6	4	100333.865	148722.349	129667.781	184814.916	60834.791
7	5	107491.311	100032.864	146640.236	117219.674	73556.337
8	6	75797.352	107168.837	98632.404	132562.773	46653.43
9	7	77763.574	75569.96	105668.473	89163.693	52759.984
10	8	56960.422	77530.283	74511.981	95524.3	35487.15
11	9	56443.131	56789.541	76444.859	67358.83	38018.671
12	10	42630.405	56273.801	55994.487	69106.153	26808.815
13	11	41080.604	42502.514	55485.968	50619.017	27504.249
14	12	31803.877	40957.362	41907.479	50159.315	20146.369
15	13	29967.283	31708.466	40383.959	37884.361	19963.407
16	14	23667.406	29877.381	31264.547	36507.099	15077.976
17	15	21901.237	23596.404	29459.098	28263.151	14529.825
18	16	17577.593	21835.533	23266.054	26631.024	11248.734
19	17	16030.787	17524.86	21529.836	21032.513	10599.148
20	18	13034.187	15982.694	17279.512	19462.972	8370.94
21	19	11748.56	12995.084	15758.937	15620.679	7746.263
22	20	9653.028	11713.314	12813.153	14246.079	6217.03
23	21	8619.018	9624.069	11549.328	11583.09	5669.939
24	22	7141.808	8593.161	9489.332	10440.593	4610.07
25	23	6328.369	7120.383	8472.857	8578.356	4155.356
26	24	5279.645	6309.384	7020.697	7659.463	3414.186

Figure 4. Simulation of the new model with GeoGebra, time diagrams for age groups A and E

6 REMARKS

The lesson could now be continued with questions like: Which value for parameter f_A would guarantee the survival of the Viennese population (if all the other parameters are kept constant)? The same could be done for parameter f_B. Are there periods of time where the age structures of the population differ considerably? What consequences does this have for a city like Vienna? And so on.

It is questions like these that make demographic change one of the greatest challenges for politics and society in developed countries (Lesthaeghe, 2014). The answers have consequences for almost all areas of life: retirement plans, child care, medical schemes, the job market, and so on. One of the educational goals of schools should be to give students an insight into this huge and important field of interest. Mathematics education can contribute to this task by simulating different scenarios of population development. GeoGebra is an ideal tool for implementing and evaluating systems of difference equations. Mathematical activities like iterative thinking, setting up systems of equations, interpreting the terms involved, using and understanding the given data and comparing it to the computed results are key elements in the presented unit (Wittmann, 2001).

References

Ableitinger, C. (2013). Bevölkerungsentwicklung mit Leslie-Modellen. *Mathe vernetzt*, *3*, 113–122.

Britton, N. F. (2012). *Essential mathematical biology*. Springer Science & Business Media.

Haspekian, M. (2005). An “instrumental approach” to study the integration of a computer tool into mathematics teaching: the case of spreadsheets. *International journal of computers for mathematical learning*, *10*(2), 109–141.

Lesthaeghe, R. (2014). The second demographic transition: A concise overview of its development. *Proceedings of the National Academy of Sciences*, *111*(51), 18112–18115.

Statistik Austria. (2017). Armut und soziale Eingliederung. *Poverty and social inclusion. Retrieved March*, *16*, 2017.

Weigand, H. G. (2004). Sequences - basic elements for discrete mathematics. *ZDM*, *36*(3), 91–97.

Wittmann, G. (2001). Tabellenkalkulation und diskrete Mathematik in der Sekundarstufe I. *Mathematikunterricht*, *47*(3), 38–51.

Christopher Ableitinger, is a mathematics faculty member at Center for Teacher Education at the University of Vienna. His professional interests include didactics of mathematics, focusing on transitions between school and university, research and development of teacher competences, use of videos in university teaching, and application-oriented mathematics education including biomathematics, game theory, and modeling.