Mathematics Analysis and Approaches Higher Level Internal Assessment

Session May 2021

Modelling population sizes and changes

1. Introduction

In this IA I will be analysing the effects of introducing a predatory species to an environment. A real life scenario where such a case took place is Yellowstone national park in the United States of America. Here wolves, previously a part of the ecosystem, were hunted to extinction. This absence of a predator gave Yellowstone's elk the conditions they needed to explode in population size. The explosion caused numerous problems in the ecosystem since the elk would eat the forest and plants faster than they could regrow, effectively destroying the habitat of countless animals. After the reintroduction of wolves the elk population dropped drastically as a direct consequence.[1]

I'm in general interested in the data driven modern world where companies like Google effectively make their profit off of data. Animal populations are also close to my heart. Living basically in a forest, my area of habitat has seen a significant increase in deer population - an animal not too different from elk - almost becoming a problem. Here the local officials are trying to keep the deer population under control by issuing larger hunting quotas, while in Yellowstone they introduced wolves to the environment. One could say that following events similar to Yellowstone's chain of events in person brought me closer to the topic.

The purpose of this investigation is modeling the general trends of population changes due to the introduction of a predatory species. I decided to construct two models of the population. One would be a time series simulation where the change in population is calculated for every year based on the previous years' population and the other an equation based on the data available and part of the aforementioned model. The prediction could be used to model effects of introducing predatory species to curb population growth of other species, such as deer, which have grown to cause a burden on the environment to prevent the "cobra effect", tampering with the environment to fix a problem and making it worse.

[1] Smith, 2016

2.Calculations

2.1 Raw Data

Figures 1 and 2 depict the data of the populations of wolves and elk respectively in Yellowstone park. Since access to the sources for these graphs couldn't be secured, I estimated the data from the images to tables 1 and 2. This makes the analysis slightly less accurate, however the data is already mathematically inconsistent due to nature being unpredictable and inconsistent in the short term causing inexplicable fluctuations - it can be considered a random sampling of the underlying distribution of elk. The focus of the models is on long term developments, which this data allows. For some years the data for the elk population was not available in figure 2 and hence is missing from tables 1 and 2.

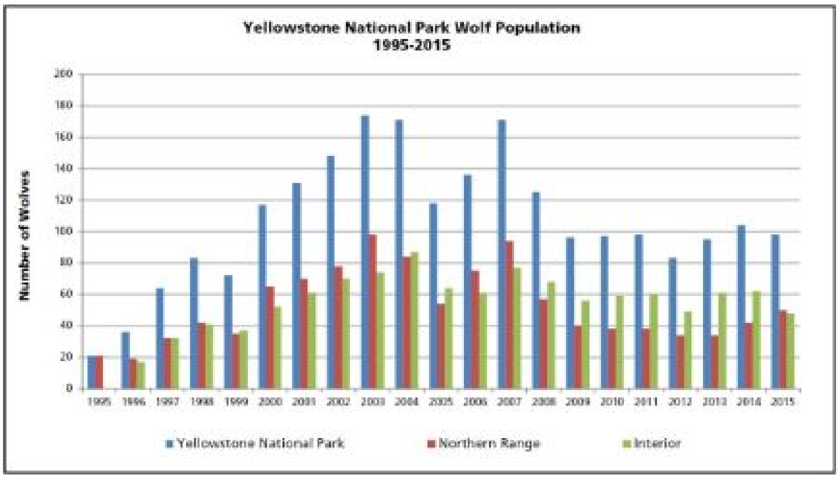

Figure 1: Wolf population data[2]

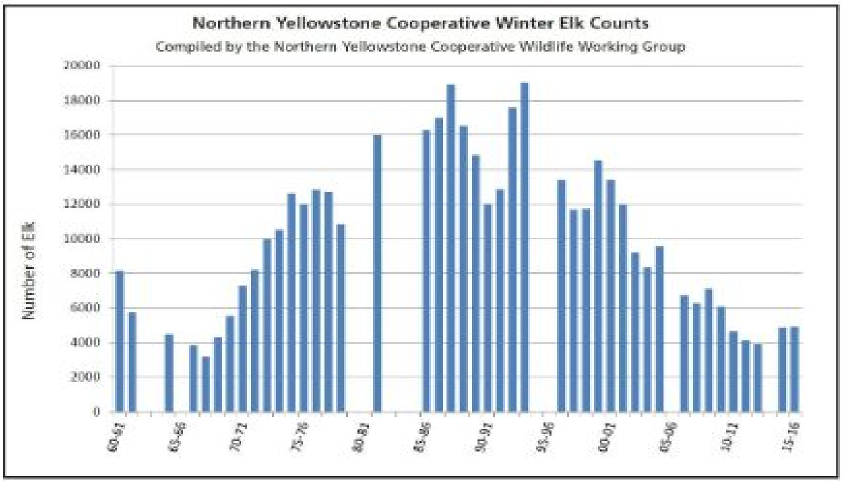

Figure 2: Elk population data[3]

[2] Smith, 2016
[3] Smith, 2016

Table 1: Extracted data pre 1995[4]

Elk population 1967 to 1995	
Year	Elk population
1967	**3000**
1968	**4100**
1969	**5200**
1970	**7000**
1971	**8100**
1972	**10000**
1973	10200
1974	12300
1975	12000
1976	12400
1977	12500
1978	10500
1979	-
1980	-
1981	16000
1982	-
1983	-
1984	-
1985	16200
1986	17000
1987	18800
1988	16400
1989	14800
1990	12000
1991	13000
1992	17500
1993	19000
1994	-

Table 2: Extracted data post 1995

Elk and wolf population 1995 to 2015		
Year	Elk population	Wolf population
1995		**40**
1996	13400	**72**
1997	11800	**122**
1998	11900	**163**
1999	14500	145
2000	13000	230
2001	12000	260
2002	9000	296
2003	8200	345
2004	9500	337
2005	-	232
2006	6500	270
2007	6200	343
2008	7000	248
2009	6000	192
2010	4300	195
2011	4100	196
2012	3900	164
2013	-	192
2014	5000	207
2015	5000	193

2.2 Simulation using Time series

This simulation method calculates each year’s elk and wolf population from the previous years’ populations via a formula that depends on earlier values. I’ll first develop the formula for the elks alone followed by a formula for the wolves and their impact on the elk population.

[4] Smith, 2016

2.2.1 Elk's equation

First an equation for the elk population. The time series will be denoted as E_n and W_n for the n^{th} year after 1965 for elk and wolves respectively. The equation the calculation is based off of is a modified version of the logistic equation from the course book[5]:

$$\frac{dP}{dt} = R \times (1 - \frac{P}{M}) \times P_{n-1}$$

$$P_n = P_{n-1} + R \times (1 - \frac{P}{M}) \times P_{n-1}$$

R = birth rate, denoted as R_E and R_W for elk and wolves respectively.

M = maximum population environment can sustain, M_E for elk and M_W for wolves.

The equation for the elk population will include a deduction of the elks killed by wolves each year which will simply be the number of wolves multiplied by the number of elks a wolf kills each year. Since a wolf typically kills 16-22 elk annually[6] and the equation needs a stable number, we take the number of elk killed by a single wolf in a year as 19, the average of the range. Hence the number of elk killed by wolves at a given year is D=19. Here W_{n-1} denotes the wolf population of the previous year. Now the equation looks like this:

$$E_n = R_{n-1} + R_E \times (1 - \frac{E_{n-1}}{M_E}) \times E_{n-1} - D \times W_{n-1}$$

Now I will add a mechanic to factor in the elk dying of causes excluding wolves. In nature animals rarely live to their full life expectancy due to predators and resources and hence a simple mechanic of removing the animals born when they reach the end of their lifespan seems inadequate. The best option is distributing these deaths over the lifespan of the animals. It would follow logic that the distribution would lean toward the end of the animal's lifespans, but a linear distribution also works since it corrects itself with the following generations also dying prematurely, which balances out in the long run. In this way the constant imbalance balances in the bigger picture and following the principle of modeling the general direction. Hence the death rate would be the average of the sum of new elk of the previous L_E years, where L_E is the average lifespan of elk in Yellowstone park.

[5] Wazir, 2019, pp.849
[6] Thuemberg, 2017

$$\frac{1}{L_E} \times \sum_{k=0}^{L_E} (E_{n-k} - E_{n-k-1})$$

Placing this to the time series equation gives us the following.

$$E_n = E_{n-1} + R_E \times (1 - \frac{E_{n-1}}{M_E}) \times E_{n-1} - D \times W_{n-1} - \frac{1}{L_E} \times \sum_{k=0}^{L_E} (E_{n-k} - E_{n-k-1})$$

Next I will estimate the maximum elk population. This can be done based on the data in table 1. Since the elk were already an established species in 1967, before the introduction of the wolves, the data, instead of a logistic curve, resembles an inverse exponent curve. LoggerPro[7] offers an automatic curve fit optimisation option to a given set of data, provided that the user provides LoggerPro with the equation. To best fit it the data was first standardised, so moved to the left by 1954 to center it according to the y axis and then divided by 20 000 to compress the data to the range 0 < x < 1. Here f(x) is the value in the new graph and p(x) is the previous population for a given year. So each datapoint goes through the transformation: $f(x) = \frac{p(x - 1954)}{20\,000}$. Fitting an inverse exponent curve as a curved best fit on the standardized data gives the following graph:

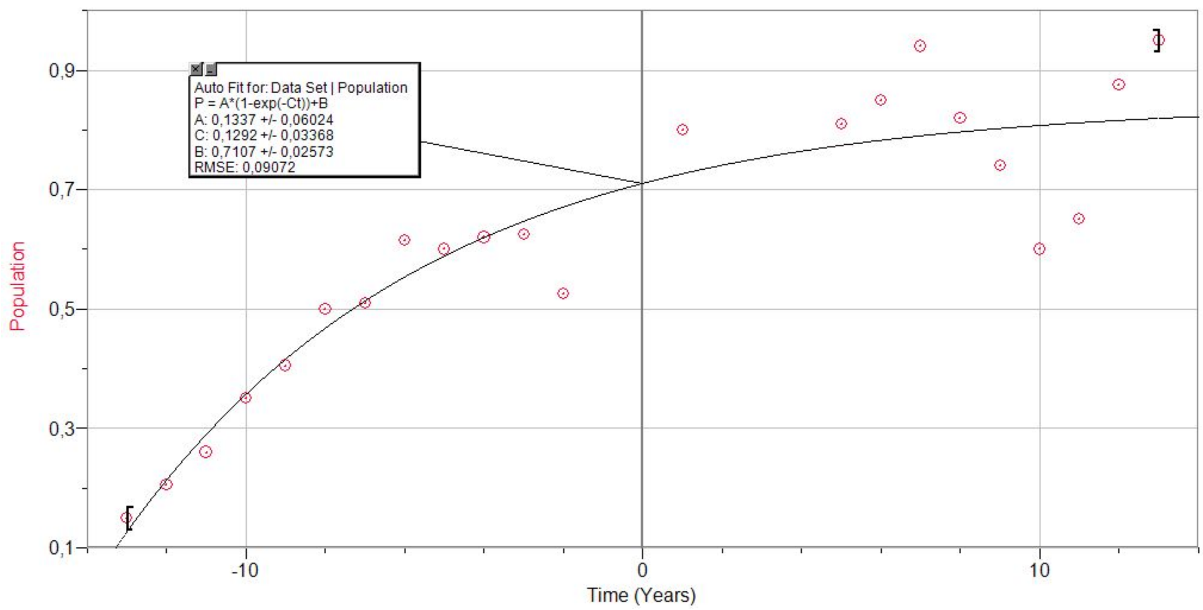

Figure 3: Curve fit for maximum population, modelled with Logger Pro[8]

[7] LoggerPro
[8] "Logger Pro."

This graph has the equation $y = 0.1337 \times (1 - e^{-0.1292x}) + 0.07107$, resembling the elk population per year, divided by 20 000. To find E_M I'll find the limit of this equation and multiply it by 20 000 to reverse the standardisation of the data. Since x will equal infinity there is no point in adding 1954 back to x.

$$M_E = 20000 \times \lim_{x\to\infty} (0.1337 \times (1 - e^{-0.1292x}) + 0.07107) = 20000 \times (0.1337 - \frac{0.1337}{e^{0.1292\times\infty}} + 0.07107)$$
$$M_E = 20000 \times (0.1337 - 0 + 0.07107) = 20000 \times 0.8444 = 16888$$

The lifespan of elk in yellowstone is about 13 years[9]. Since $E_1 = 3000$, from table 1, the only remaining variable is R_E. Because the investigation intends to analyse the effect of the introduction of the predatory species and to make a prediction, R_E can simply be found for the data in table 1 and figure 4.

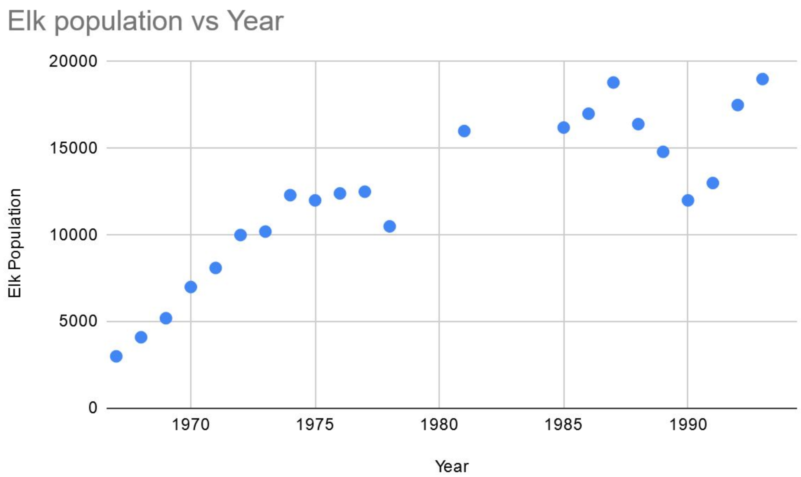

Figure 4: Elk per year, modelled with Google Sheets[10]

The first six data points are most consistent and will hence be used for the calculation, since in general in nature fluctuations in the population happen quite a bit. The points are highlighted in Table 1 corresponding to years 1967-1972. Putting those values into the death per year formula returns approximately 623 deaths in 1972 from non-wolf related causes. The wolf factor, $19W_{n-1}$, is not present yet since the wolves have not been introduced yet.

$$E_n = E_{n-1} + R_E \times (1 - \frac{E_{n-1}}{16888}) \times E_{n-1} - N_{Dead}$$

$$10000 = 8100 + R_E \times (1 - \frac{8100}{16888}) \times 8100 - 623.15...$$

[9] "Elk", 2020
[10] "Sheets."

$$R_E = \frac{10000-8100+623.15...}{8100\times(1-\frac{8100}{16888})} = \frac{2523.15...}{4214.88...} = 0.598... \approx 0.599$$

Now we have an equation for E_n from E_{n-1}, the Elk model is ready.

$$E_n = E_{n-1} + R_E \times (1 - \frac{E_{n-1}}{M_E}) \times E_{n-1} - D \times W_{n-1} - \frac{1}{L_E} \times \sum_{k=0}^{L_E} (E_{n-k} - E_{n-k-1})$$

$$\Rightarrow E_n = E_{n-1} + 0.599 \times (1 - \frac{E_{n-1}}{16888}) \times E_{n-1} - 19 \times W_{n-1} - \frac{1}{13} \times \sum_{k=0}^{13} (E_{n-k} - E_{n-k-1})$$

Using this equation with E_1 = 3000 we get the following model for the years 1967-1994:

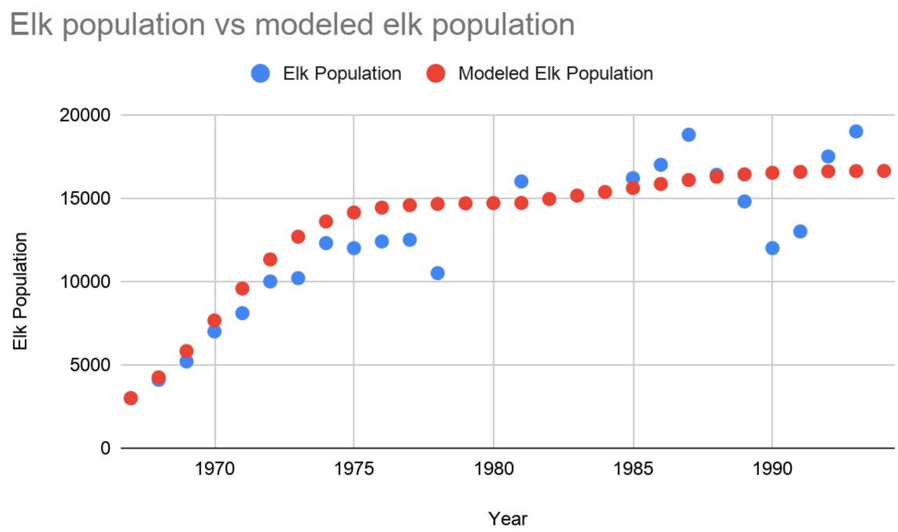

Image 4: Elk model comparison, modelled with Google Sheets[11]

The modelled data clearly follows the actual data from Yellowstone park. Now I'll move on to create a model for the wolves.

2.2.2 Wolves' equation

Now a somewhat similar time series equation will be constructed for the wolves of Yellowstone park. The main difference is that the maximum population won't be fixed and there won't be a factor for deaths by falling prey to wolves. W_1 = 40 from Table 1, L_W is the lifespan of wolves, M_W is the maximum population and R_W is the rate of reproduction.

$$W_n = W_{n-1} + R_W \times (1 - \frac{W_{n-1}}{M_W}) \times W_{n-1} - \frac{1}{L_W} \times \sum_{k=0}^{L_W} (W_{n-k} - W_{n-k-1})$$

[11] "Sheets."

The average lifespan of a wolf in Yellowstone park is 4 to 5 years[12], rounding the average to the nearest integer, to fit the sum with a whole number, $L_W = 5$. The same principles apply in the death rate of the wolves as do with the elks with the deaths being distributed over the generations lifespan. The maximum wolf population has to be relative to the current elk population, since elk are the primary prey of the wolves and there is clear causation between the wolf populations growth and the elk populations diminishment[13]. Since the average wolf consumes 19 elk a year we can make an assumption on the total number of elk per wolves required for the wolves to reach this elf consumption. The wolves will not catch every elk in the park. A reasonable estimate for the required elk count would be twice the amount of elk caught, 38 elk per wolf, if the wolves are capable of hunting at most half the elk in the park in a given year due to the sheer size of the park - it's effectively impossible for them to cover all the area and find all the elk. Hence $M_E = \frac{E_{n-1}}{38}$ and the full equation that follows would be:

$$W_n = W_{n-1} + R_W \times (1 - \frac{W_{n-1}}{\frac{E_{n-1}}{38}}) \times W_{n-1} - \frac{1}{L_W} \times \sum_{k=0}^{L_W} (W_{n-k} - W_{n-k-1})$$

This again leaves us with R_W, the reproduction rate of the wolves, to be found. This can be done the same way as it was previously, by substituting values from the live data into the equation. With the wolves the first four values fill the same criteria as the first six with the elks - they're consistent. The death count for this time period adds up to 24.4, giving us the following calculation.

$$W_n = W_{n-1} + R_W \times (1 - \frac{W_{n-1}}{\frac{E_{n-1}}{38}}) \times W_{n-1} - N_{Dead}$$

$$163 = 122 + R_W \times (1 - \frac{122}{\frac{11800}{38}}) \times 122 - 24.4$$

$$R_W = \frac{163+24.4-122}{122\times(1-\frac{122\times38}{11800})} = \frac{65.4}{74.06...} = 0.582... \approx 0.883$$

Fitting in the values we have

$$W_n = W_{n-1} + 0.883 \times (1 - \frac{W_{n-1}}{\frac{E_{n-1}}{38}}) \times W_{n-1} - \frac{1}{5} \times \sum_{k=0}^{5} (W_{n-k} - W_{n-k-1})$$

[12] "Gray Wolf", 2020
[13] Smith, 2016

2.2.3 Compiling the model

The equations for the two populations now are the following for elks and wolves respectively.

$$E_n = E_{n-1} + 0.599 \times (1 - \frac{E_{n-1}}{16888}) \times E_{n-1} - 19 \times W_{n-1} - \frac{1}{13} \times \sum_{k=0}^{13} (E_{n-k} - E_{n-k-1})$$

$$W_n = W_{n-1} + 0.883 \times (1 - \frac{W_{n-1}}{\frac{E_{n-1}}{38}}) \times W_{n-1} - \frac{1}{5} \times \sum_{k=0}^{5} (W_{n-k} - W_{n-k-1})$$

Table 3: Original populations vs modelled populations 1995-2015

	Original Data Elk vs Wolves		Modelled Elk vs Wolves	
Year	Elk	Wolf	Elk model	Wolf model
1995		40	16634	40
1996	13400	72	15876	64
1997	11800	122	15156	99
1998	11900	163	14202	145
1999	14500	145	12886	194
2000	13000	230	11227	229
2001	12000	260	9488	237
2002	9000	296	7988	213
2003	8200	345	7098	188
2004	9500	337	6712	178
2005		232	6498	180
2006	6500	270	6245	181
2007	6200	343	5952	176
2008	7000	248	5739	164
2009	6000	192	5724	156
2010	4300	195	5860	156
2011	4100	196	5966	159
2012	3900	164	5956	162
2013		192	5827	160
2014	5000	207	5614	155
2015	5000	193	5350	149

Graphing table 3 gives figure 5.

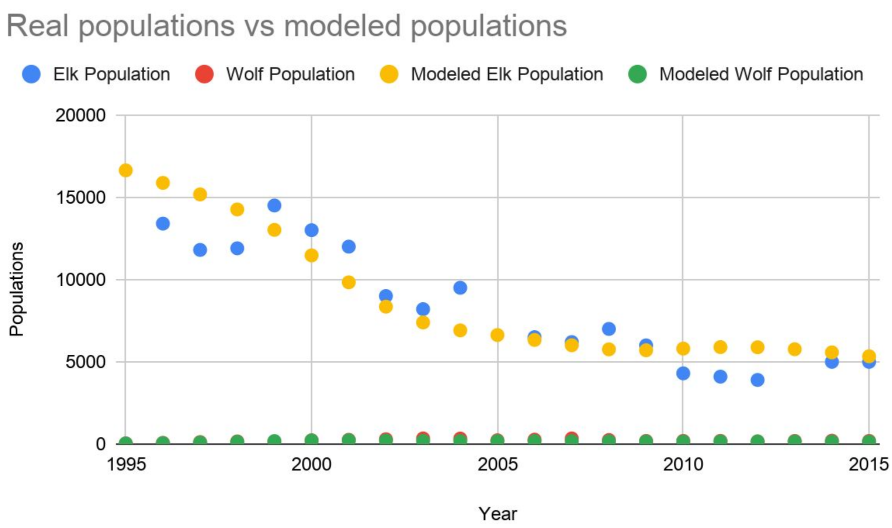

Figure 5: Simulation vs reality, modelled with Google Sheets[14]

The elk's model follows the elk population quite closely, clearly following the larger trends. The model has less fluctuation, but this is to be expected since in nature there are countless factors in play affecting the annual population of either animal. In order to better assess the accuracy of the wolf model a separate graph was drawn in figure 6.

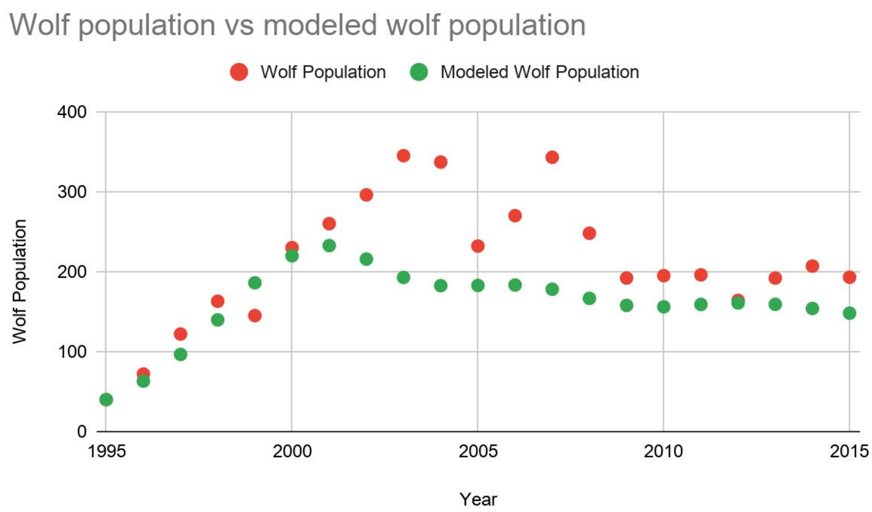

Figure 6: Wolf model assessment, modelled with Google Sheets[15]

From figure 6 it is clear that the wolf model deviates from reality in having a lower peak and a bit lower of a level after 2008 in general. Yet the model quite closely resembles reality. Hence the relationship between the models is functional and the simulation works.

---

[14] "Sheets."
[15] "Sheets."

The mean error of the simulation for the elk, 1967 - 2015, can be calculated by taking the sum of the difference of each modeled data point and the original data and dividing this sum by the number of datapoints. The calculation can only be completed for the years for which we have the original data from tables 1 and 2. This means excluding years 1979, 1980, 1982, 1983, 1984, 1994, 1995, 2005 and 2013 leaving us with 40 data points. I will compare this to the accuracy of the equation I'll construct later.

$$Total\ error = \sum_{t=1}^{40} (P_{Data} - P_{Model})$$

$$Mean\ error = \frac{1}{40} \times \sum_{t=1}^{40} (P_{Data} - P_{Model})$$

Total error = 22243, Mean error = 556

Even though 556 elk might seem large at first, it's only 3.2% of the max elk population without wolves - though it is 19% of the maximum with the wolves. I'll address this uncertainty further in the conclusion. It's important to choose the right error metric for the specific case. In the specific context of the investigation the mean error is preferable to mean absolute error since the simulation aims to find the larger change, the average. So if the values go under but fix by going over, then the general direction is correct. This is something which would be punished in mean absolute error calculations but not when using mean error. Mean square error similarly punishes the system for missing fluctuations, which isn't the goal of this simulation.

2.2.4 Modelling the model

Next I'll develop a model for the system post 1995. This portion will model the elk population, since it is the more significant of the two to the people operating Yellowstone national park - the whole operation is after all to curb the elk population while wolves are actively being hunted[16]. Constructing an equation to fit the elk data is extremely difficult since the fluctuation between years isn't too orderly. If we disregard the fluctuation we can optimise a

[16] Thuemberg, 2017

natural exponent equation to fit our needs "$P = A \times e^{-C \times t} + D$". Here P is the elk population t years after 1965. D is the number it stabilises around and can be found by taking the average of the last 15 terms, which is approximately 2783. A in the equation is the furthest distance from D which the graph reaches, or the starting population - D. A is just the difference between the initial value and D. $A = 16634.6 - 2783.3 = 13850.2$. C can be isolated and solved for as follows. The point used for the calculation should be as close as possible to the average of the peak and average value in order to acquire the most accurate graph. The average is 9708, which is closest to the sixth data point but to be sure the calculation was completed with both the sixth and the fifth point and the RMSE, Root Mean Square Error, of the fifth datapoint was 546.509 units while the sixth gave 588.689, as calculated with LoggerPro and as is demonstrated in figure 8. Hence the point (5, 11227) was used. Results can be seen in figure 7.

$$P = A \times e^{-C \times t} + D$$
$$C = \frac{-1}{t} \times ln\left(\frac{P-D}{A}\right) = \frac{-1}{5} \times ln\left(\frac{11227-2783}{13850}\right) = 0.09898... \approx 0.09898$$
$$P = 13850 \times e^{-0.9898 \times t} + 2783$$

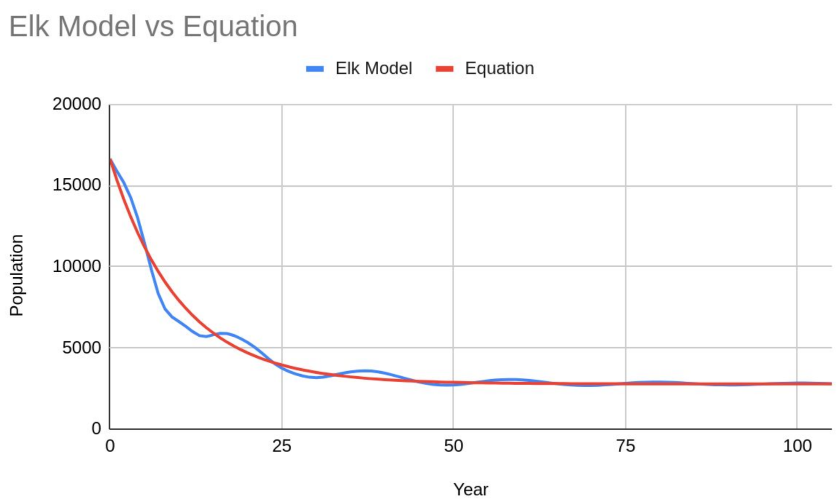

Figure 7: Red line as the equation for P, modelled with Google Sheets[17]

[17] "Sheets."

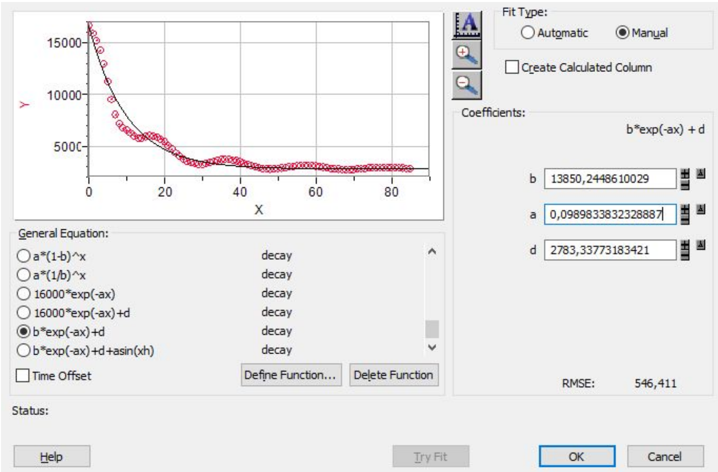

Figure 8: RMSE calculation from Logger Pro[18]

This equation alone is a very rough estimate of what the elk population will look like x years after 1995. A different approach to this problem model will utilise this equation. For this I will first format an equation for the elk population without the wolves and then adjust for wolves come 1995 - or any other year for analysis purposes. The benefit of such a formula is that we can see the entire situation starting all the way from 1967. For this purpose we will deviate from the very beginning in how the equation is formatted.

2.3 Fitting an Equation on the Data

For the equation we will deviate from the very beginning. The equation will still be based on the logistic equation from the course book[19], similarly to the time-series method:

$$\frac{dP}{dt} = R \times (1 - \frac{P}{M}) \times P_{n-1}$$

R = rate of reproduction

L = 16888, maximum population of Elk

This time we take the change in population and integrate it to find an equation for the whole population. The integration is as follows:

$$\frac{dP}{dt} = R \times P \times (1 - \frac{P}{M})$$

$$dP = (P - \frac{P^2}{M}) \times R \times dt$$

$$R\,dt = (P - \frac{P^2}{M})^{-1}\,dP$$

[18] “Logger Pro.”

[19] Wazir, 2019

$$(P - \frac{P^2}{M})^{-1} = \frac{A}{P} + \frac{B}{1 - P \times M^{-1}} = \frac{A - A \times P \times M^{-1} + B \times P}{P \times (1 - P \times M^{-1})}$$

$$\Rightarrow A = 1 \;\; and \;\; (B - A \times M^{-1}) = 0 = (B - M^{-1}) \;\Rightarrow B = M^{-1}$$

$$\Rightarrow (P - \frac{P^2}{M})^{-1} = \frac{1}{P} + \frac{M^{-1}}{1 - P \times M^{-1}} = \frac{1}{P} + \frac{1}{M - P}$$

$$\Rightarrow R\, dt = \left(\frac{1}{P} + \frac{1}{M - P}\right) dP$$

$$\Rightarrow \int R\, dt = \int\left(\frac{1}{P} + \frac{1}{M - P}\right) dP = \int \frac{1}{P}\, dP + \int \frac{1}{M-P}\, dP$$

$$R \times t + c = ln\,|P| - ln\,|M - P| = ln\left|\frac{P}{M-P}\right| \Rightarrow \frac{P}{M-P} = e^{R \times t + c}$$

$$P = (M - P) \times e^{R \times t + c} = M \times e^{R \times t + c} - P \times e^{R \times t + c}$$

$$M \times e^{R \times t + c} = P + P \times e^{R \times t + c} = P \times (1 - e^{R \times t + c})$$

$$P = \frac{M \times e^{R \times t + c}}{1 + e^{R \times t + c}} = \frac{M \times e^{R \times t + c}}{1 + e^{R \times t + c}} \times \frac{e^{-(R \times t + c)}}{e^{-(R \times t + c)}} = \frac{M}{e^{-R \times t - c} + 1} = \frac{M}{1 + C \times e^{-R \times t}}$$

A similar task can also be found in the coursebook[20], without the integration steps. M is the maximum population and it remains the same as before at 16888. This leaves C and R. When t = 0 the population is at 3000, which lets us solve C. Next it is as simple as substituting some values from the existing data to find R.

$$M = P + P \times e^{-R \times t} \times C \Rightarrow C = \frac{M-P}{P \times e^{-R \times t}} = \frac{16888 - 3000}{3000 \times e^{-R \times 0}} = \frac{1736}{375} = 4.629\overline{3} \approx 4.63$$

$$R = \frac{-1}{t} \times ln\left|\frac{M-P}{P \times C}\right| = \frac{-1}{5} \times ln\left|\frac{16888 - 10000}{10000 \times 4.62...}\right| = 0.381... \approx 0.381$$

$$P = \frac{16888}{1 + 4.63 \times e^{-0.381 \times t}}$$

Graphing these in figure 9 we can clearly see that the formula works accurately from 1967 to 1994.

[20] Wazir, 2019, pp.850

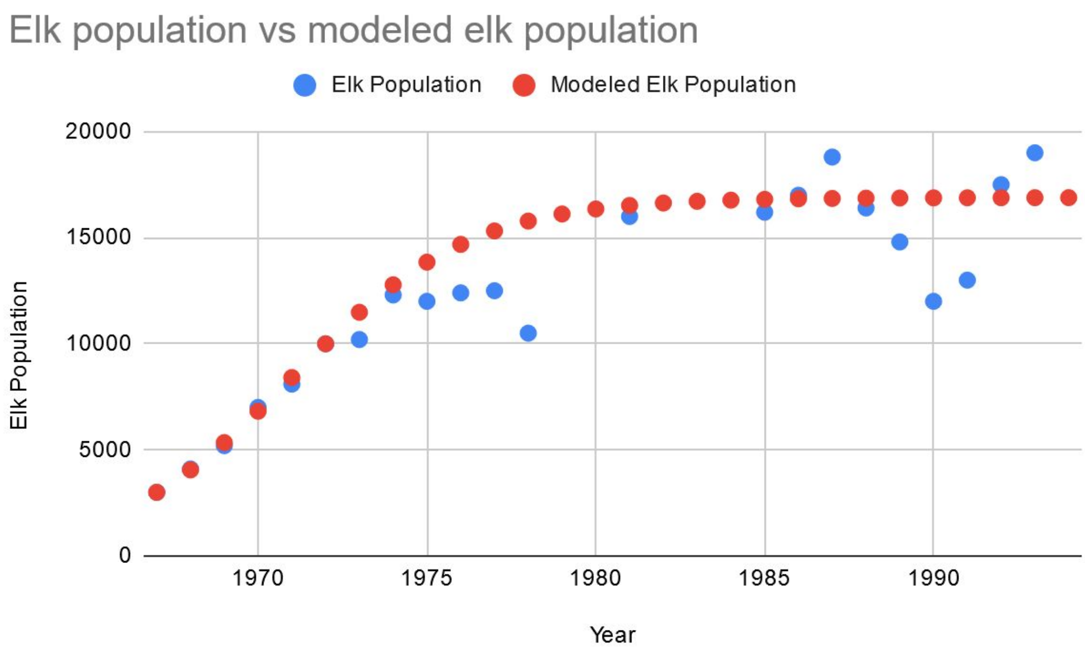

Figure 9: Population vs Year, testing the model, modelled with Google Sheets[21]

Now to continue this, we don't need to make a wolf model with a similar logic. Instead we can simply use a varying maximum population of elks with $P = 13850 \times e^{-0.9898 \times x} + 2783$ working to our advantage. Using this equation as the maximum elk population post 1995 we can easily adjust for the introduction of the wolves. In order to get all of this into a single equation we must take some steps. We will apply a sigmoid function to both values with an extremely steep slope and move P to the right by 28. The sigmoid functions will compensate for each other for the transition period, effectively always adding up to the correct number, sample sigmoid graphs are displayed in figure 10.

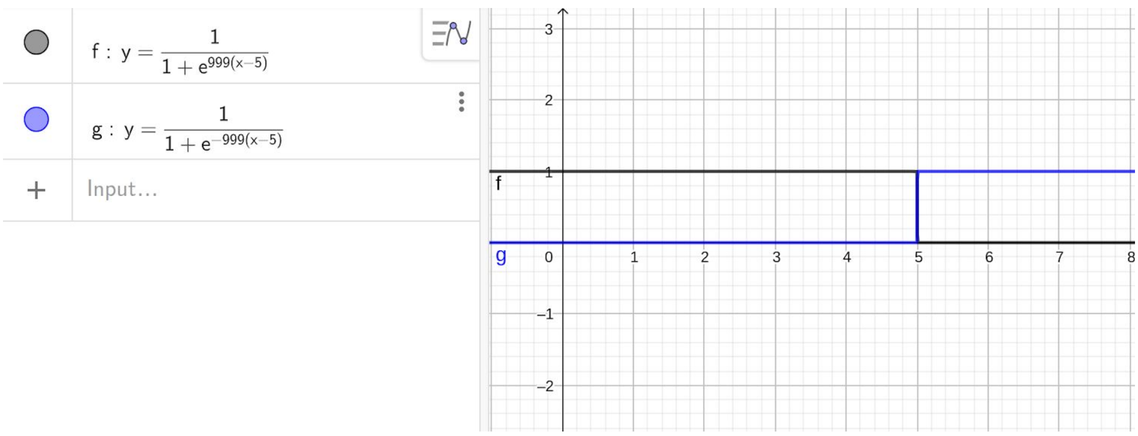

Figure 10: Sigmoid functions, modelled with GeoGebra[22]

[21] "Sheets."
[22] Hohenwarter, Geogebra

The steps will result in the following equation P for population t as years after 1967.

$$\sigma_1 = \frac{1}{1 + e^{(-999 \times (t - 28))}}, \quad \sigma_2 = \frac{1}{1 + e^{(999 \times (t - 28))}}$$

$$p = 13850 \times e^{-0.9898 \times x} + 2783, \quad M = 16888$$

$$P = \frac{\sigma_1 \times p + \sigma_2 \times M}{1 + 4.63 \times e^{-0.81 \times t}}$$

$$P = \frac{(13850 \times e^{-0.9808 \times (t - 28)} + 2783) \times \frac{1}{1 + e^{(-999 \times (t - 28))}} + 16888 \times \frac{1}{1 + e^{(999 \times (t - 28))}}}{1 + 4.63 \times e^{-0.81 \times t}}$$

$$P = \frac{\frac{13850 \times e^{(-0.9808 \times (t - 28))} + 2783}{1 + e^{(-999 \times (t - 28))}} + \frac{16888}{1 + e^{(999 \times (t - 28))}}}{1 + 4.63 \times e^{-0.81 \times t}}$$

This equation can be seen in action in figure 11.

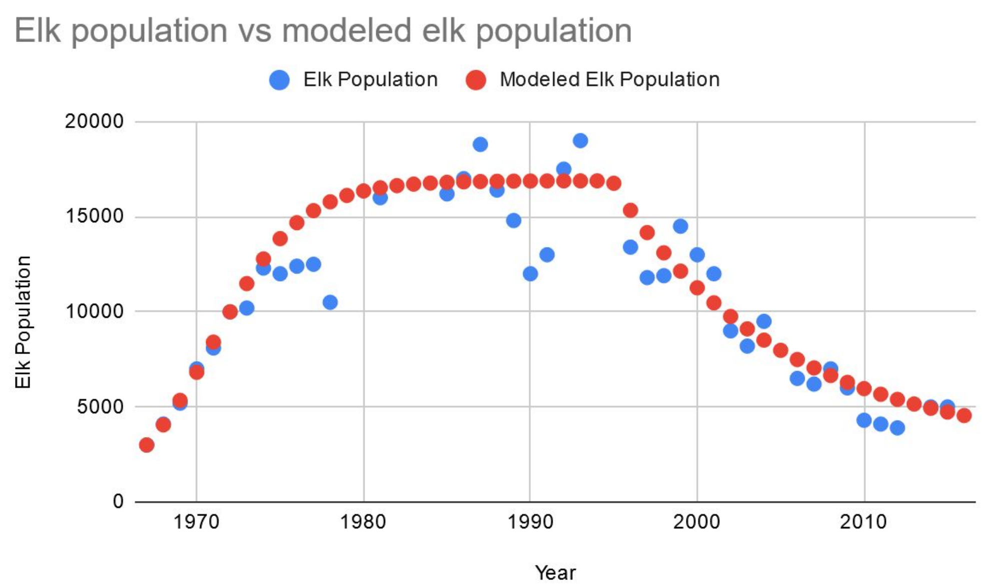

Figure 11: final equation in action, modelled with Google Sheets[23]

It is clear that the model follows the data, practically as accurately as a model can when it has been constructed based on data that was acquired from visual analysis of an image. In essence the model shows the larger shifts and developments in the elk population before and after the introduction of wolves. Completing the same mean error calculation for the equation as the simulation shows that the equation is a bit less accurate than the simulation.

$$Total\ error = \sum_{t=1}^{40} (P_{Data} - P_{Model})$$

[23] "Sheets."

$$Mean\ error = \frac{1}{40} \times \sum_{t=1}^{40} (P_{Data} - P_{Model})$$

Total error of equation = 28514, Mean error of equation = 713

Total error of simulation = 22243, Mean error of simulation = 556

2.4 Application

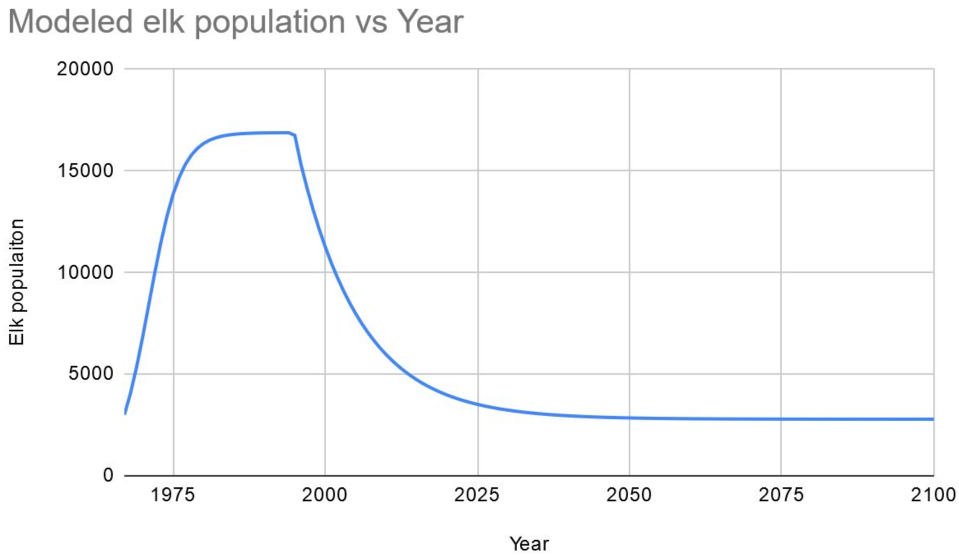

Figure 12: graphing until 2100, modelled with Google Sheets[24]

With the equation it is very easy to predict the elk population for the foreseeable future. While this equation does lack the fluctuation effect of the calculation model, it is considerably easier to use and clearer for any observatory purposes. With an equation like this the officials in charge of determining hunting quotas for wolves and elk could clearly see whether it is necessary. Unless dramatic unforeseeable change overtakes the Yellowstone national park ecosystem, increasing the wolf hunting quota should not be necessary. However even if it is increased, in the worst case as long as 40 wolves remain we can be sure that the wolf population will regrow since the original introduction included 40 wolves. By adjusting the formula mildly we can see how introducing the wolves at different points between 1967 and 1995 would affect the curve. This is simply done by adjusting the sigmoid equations and the elk max population equation toward the left - or less toward the right from where they

[24] "Sheets."

currently are. This can be seen in practise in figure 13 in a model of the elk population where wolves were introduced in 1972 instead of 1995.

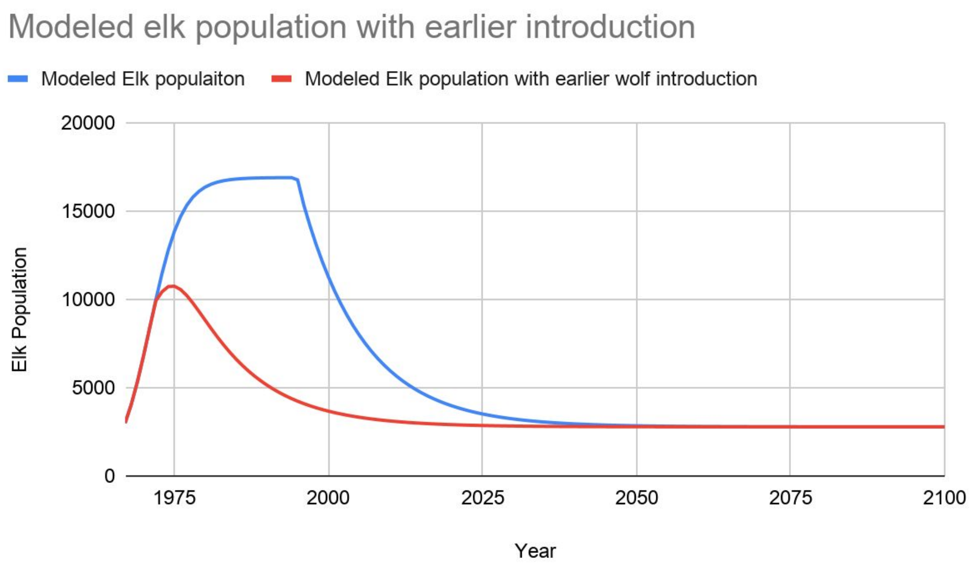

Figure 13: Elk population if wolves were introduced in 1972, modelled with Google Sheets[25]

The earlier introduction of the wolves caused a drop in elk population to the expected level, which is the level where the wolf and elk populations balance out at. In reality there would be a slight fluctuation around the balance position since as the wolf population grows it'll hunt more elk which leads to a lack of elk and starvation the next year. This starvation would lead to an increase in the elk population since less wolves are hunting them. The cycle repeats. However for modelling purposes the fluctuation is not too significant since the point is to optimise averages and greater developments. Still this fluctuation is captured in the year by year calculation method, see figure 14, making it better for analysing individual years - you'll know which are the more populous years for the elk, even if the fluctuations are not perfectly precise. Again the principle of finding the general direction applies. Similarly, if desired, the year by year model can be with even more versatility to investigate hypothetical scenarios with different numbers of elk and wolves with different time periods. Even an increase in annual hunting could potentially be added to the model as a constant death factor not too different from the current elk killed by wolves - factor.

---

[25] "Sheets."

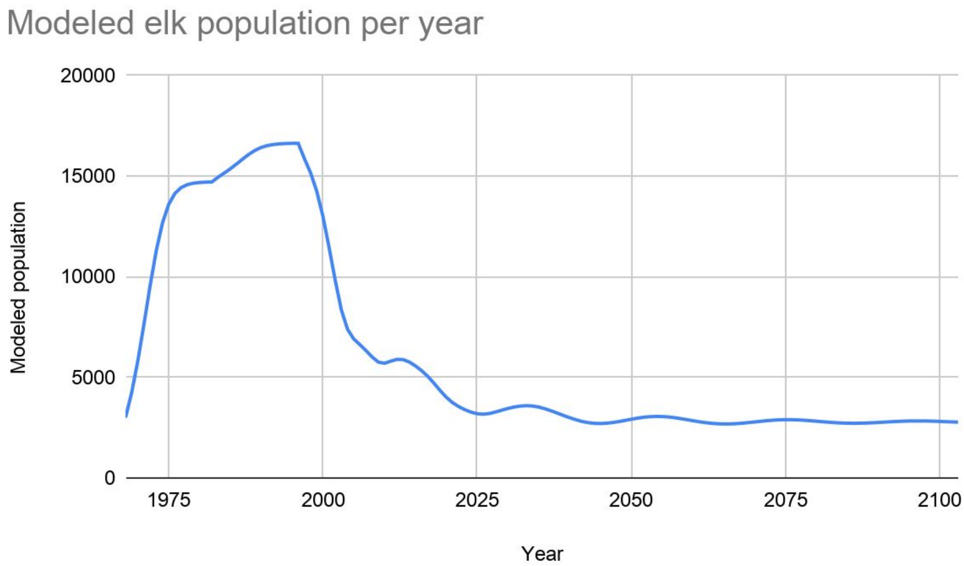

Figure 14: year by year calculation up to 2100, modelled with Google Sheets[26]

3. Conclusion

The two methods of modelling the elk population utilise one of these systems of equations.

The simulation:

$$E_n = E_{n-1} + 0.599 \times \left(1 - \frac{E_{n-1}}{16888}\right) \times E_{n-1} - 19 \times W_{n-1} - \frac{1}{13} \times \sum_{k=0}^{13} \left(E_{n-k} - E_{n-k-1}\right)$$

$$W_n = W_{n-1} + 0.883 \times \left(1 - \frac{W_{n-1}}{\frac{E_{n-1}}{38}}\right) \times W_{n-1} - \frac{1}{5} \times \sum_{k=0}^{5} \left(W_{n-k} - W_{n-k-1}\right)$$

And the equation:

$$P = \frac{\frac{13850 \times e^{(-0.9808 \times (t-28))} + 2783}{1 + e^{(-999 \times (t-28))}} + \frac{16888}{1 + e^{(999 \times (t-28))}}}{1 + 4.63 \times e^{-0.81 \times t}}$$

Both of these are valid methods and, as discussed in the “Application” portion of the investigation, have their own strengths and weaknesses. These formulas can be modified with relative ease to be applied outside Yellowstone park. They offer a way of linking two closely causation-related species’ populations together in a model where the species which is most relevant can be examined in more detail. Since data driven decision making is in my opinion usually the best kind, models such as these serve a clear purpose in managing the environment, especially a relatively closed one like Yellowstone park. While the mean error of the equation was slightly larger, the errors of both systems were on average above 500

[26] “Sheets.”

elk per year. This feels like a large number, but is explained by the random sampling of the distribution of elk being quite imperfect and fluctuating quite a bit. Additionally animal populations can fluctuate drastically in nature. The models aren't quite as good at estimating the fluctuations, causing an uncertainty, but are very competent in finding the general trends - which was my goal in this investigation. Since the purpose of the investigation was modeling the general trends of population changes due to the introduction of a predatory species the investigation is a success in achieving this goal despite the poor initial data.

An interesting extension would be adding a consideration of the fluctuations in the populations to the equation. This is very challenging since the fluctuations aren't very consistent in their changes and hence pose a challenge for someone intending to model them with a single equation. A drawback of such an addition would be the extension to an already long and complicated equation. Another extension might be adding more species to the relationship - also a weakness of this investigation is only considering two species where in nature multiple would affect the results. The wolves don't eat elk exclusively, so factoring in another prey animal could give more insight to the situation in Yellowstone. Of course access to more accurate data would benefit any investigation in the area.

Though I had heard about the application of a predatory species in Yellowstone park earlier, I didn't immediately realise its potential for the investigation I was planning. I was early on interested in modelling predator-prey interaction and the effects of predators on prey, even constructing a theoretical model which I used as a basis for the simulation in this investigation. Finding a situation where such interaction had taken place and which had distinct data turned out to be quite challenging, even if real life applications are numerous. Yet the data acquired from Yellowstone proved very much sufficient for the investigation.

Overall the investigation was a success with a working year by year model and an accurate equation displaying the development of the elk population before and after the wolves were introduced to the environment. I am personally happy with the results of the investigation.

Bibliography

"Elk." Elk - Yellowstone National Park (U.S. National Park Service), U.S. National Park Service, 2020, nps.gov/yell/learn/nature/elk.htm.

Hohenwarter, Markus. GeoGebra, 6.0.620, 2002, geogebra.org/classic#graphing.

"Gray Wolf." Gray Wolf - Yellowstone National Park (U.S. National Park Service), U.S. National Park Service, 2020, nps.gov/yell/learn/nature/wolves.htm.

"Logger Pro." Logger Pro® 3 - Vernier, 3.8.6.1, Vernier, vernier.com/product/logger-pro-3/.

"Sheets." Google Sheets, Google, docs.google.com/spreadsheets/.

Smith, Douglas W., et al. "YS 24-1 Wolf Restoration in Yellowstone Reintroduction to Recovery." YS 24-1 Wolf Restoration in Yellowstone Reintroduction to Recovery - Yellowstone National Park (U.S. National Park Service), U.S. National Park Service, 5 July 2016, nps.gov/yell/learn/ys-24-1-wolf-restoration-in-yellowstone-reintroduction-to-recovery.htm.

Thuemberg, Argus M. "How Many Elk Do Yellowstone Wolves Eat?" How Many Elk Do Yellowstone Wolves Eat?, WyoFile, 28 Mar. 2017, wyofile.com/many-elk-yellowstone-wolves-eat/.

Wazir, Ibrahim, and Tim Garry. Mathematics Analysis and Approaches for the IB Diploma. Higher Level ed., Pearson, 2019.