

Rate-based Analysis of Ecological Factors Linking Rinderpest Seroprevalence and Tree Density on the Serengeti Plain

BASEL MOSTAFA

(Dated: November 2020)

1. PROBLEM

“The migration of the wildebeests is one of the greatest spectacles on Earth. Hundreds of thousands of these animals trace a giant loop around the Serengeti each year. Interestingly, the population of wildebeests jumped from a low of roughly 400,000 in the early 1960s to over 1.4 million animals more recently. This change in the population structure was tied to the eradication of the rinderpest, a virus that was lethal to human livestock, but whose effects bled over to the wild population of wildebeests as well. Interestingly, as a result of this eradication, the number of wildebeests went up with measurable downstream changes to the entire Serengeti ecosystem. One of the most intriguing claims is that the wildebeest population increase led to a decrease in the number of wildfires resulting in turn in the increase in the number of trees and thereby larger animals such as elephants. In this hurdle, your job is to work out a model for the relation between rinderpest numbers, wildebeest numbers, fire frequencies and tree density on the Serengeti plain as well as any other factors you may think are influential in the ecosystem.”

2. ABSTRACT

This solution builds on previous research considering the correlations between rainfall, elephant population, wildebeest population, fire rates, and tree density, which establishes that the most significant causal relationships involved in this system are as depicted in 1.

This solution attempts to quantify these relationships by considering the effect of each explanatory variable on the rate of change of the corresponding response variable. Based on an understanding of these effects and some simplifying assumptions, mathematical relationships are developed to approximate the actual dynamics observed on the Serengeti. The methods used to determine these mathematical relationships combine established models for population growth (namely logistic growth), logical extensions of these models derived from simplifying assumptions, and comparison with observed and/or simulated data. The relationships between each pair of variables are then aggregated to determine the

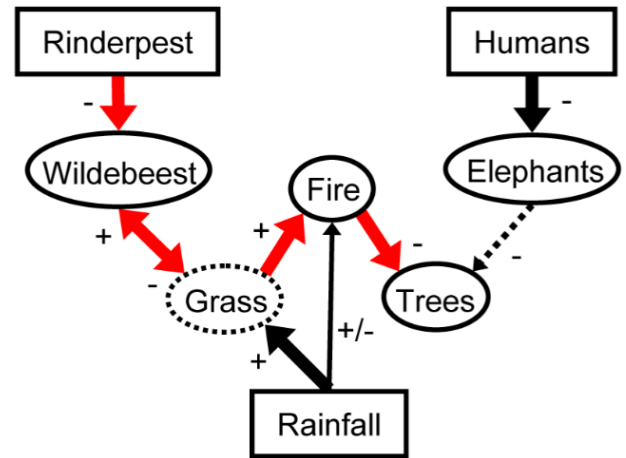


Figure 1. Primary causal relationships linking rinderpest seroprevalence to tree density on the Serengeti Plain (4).

approximated relationship between time and each of the variables considered.

3. DATA COLLECTION

The data used in this solution is provided in (4). The data was provided in table format in a word document. In order to analyze this data, it was necessary to transfer the data to a data analysis tool. The format of the data was not well-suited to this purpose, so a program was developed to format the data provided to it via a text file containing the copy-pasted table of values. This program takes a column and text file as an input, and prints the value of each row in the specified column line by line. If data is unavailable for a certain year, a blank line is printed. The result printed to the console can then be readily copy-pasted into google sheets for numerical analysis. The source code for this program is provided in the appendix.

Rinderpest data was not provided in this table. However, it was provided in a graph format. The rinderpest data was determined to two decimal places by measuring the pixel locations of each data point/labels on the axis and scaling appropriately.

4. GROWTH POPULATION MODELS

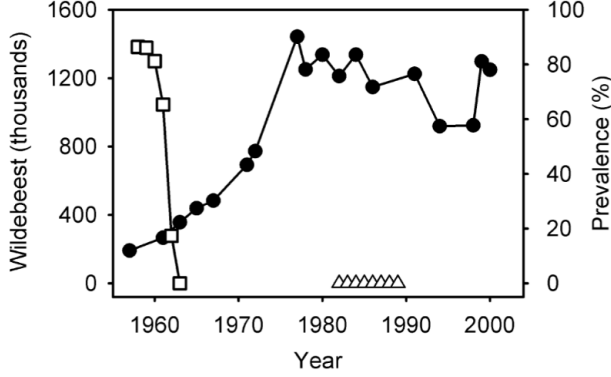


Figure 2. Wildebeest population and rinderpest seroprevalence over time. Pixel coordinates from this plot were used to determine rinderpest seroprevalence over time (4).

Population growth of organisms subject to resource limitations is often characterized by a type of density-dependent population growth model, which assumes that population growth rates are primarily dependent on two factors: the difference between the “natural” reproduction and death rates and resource limitations. These two factors combine to yield an S-shaped curve over time (7).

In general, there are three such widely-applied population growth models.

4.1. Model 1

For large populations with low birth rates (often “primitive” organisms, such as bacteria, insects, and plants), this S-shaped curve is often an exponential-logistic curve (5), characterized by the differential equation in 1, where P is population, t is time, r is the intrinsic rate of natural increase, K is the carrying capacity of the environment.

$$\frac{dP}{dt} = rP \frac{K - P}{P} \quad (1)$$

4.2. Model 2

For small populations with high birth rates (often “complex” organisms, such as larger mammals, and arguably, humans), this S-shaped curve is more strongly characterized by exponential growth until it exceeds the carrying capacity (overshoot), at which point it begins to oscillate around the carrying capacity. A purely exponential growth model of population growth is characterized by the differential equation in 2, provided in (8), where P is population, t is time, and r is the intrinsic rate of natural increase.

$$\frac{dP}{dt} = rP \quad (2)$$

However, the oscillations of population in such a population are very difficult to predict due to the concept of chaos. Chaos: “seemingly random complexity from simple deterministic equation; Not random, susceptible to initial conditions” (7).

The oscillations of a population characterized by Model 2 are extremely sensitive to the starting condition. A minuscule difference in the initial population can lead to major deviations in the population’s oscillations later on (7). This leads us to the third model.

4.3. Model 3

One can approximate Model 2 with a model that assumes exponential growth until the carrying capacity is reached, at which point the population size assumes the carrying capacity for the rest of time. This model is unlikely to exist in real life and is unsupported by evidence from any real populations, but it is often the most reasonable approximation that can be made for a population that actually follows Model 2 (5). Such a model is characterized by the differential equation in 3.

$$\frac{dP}{dt} = \begin{cases} rP & P < K \\ 0 & P \geq K \end{cases} \quad (3)$$

5. RINDERPEST SEROPREVALANCE VS WILDEBEEST POPULATION

In determining the relationship between rinderpest seroprevalence and wildebeest population, we consider the assumption that the rate of population growth is a modified form of one of the three models from Section 2. Note that the growth rate of any population can be defined as in 4, with P being population, t being time, B being births, and D being deaths.

$$\frac{dP}{dt} = \frac{dB}{dt} - \frac{dD}{dt} \quad (4)$$

Primary Assumption - We assume that the presence of rinderpest only affects the population growth rate by increasing the death rate by an amount proportional to rinderpest seroprevalence, population, and the mortality rate of the disease.

We assume that changes in rinderpest seroprevalence do not have a major impact on birth rates, even if the seroprevalence of rinderpest changes the population’s age distribution. We justify this assumption by considering the fact that if every member of a population has an equal probability of reproduction regardless of age, then an exponential growth model is intuitive. However, in the real world, the probability of reproduction depends on age, and the age distribution of a population often changes as the population increases in size

(if the death rate per capita stays constant, which is often the case, the base of a population pyramid must increase before the higher levels can grow wider). Since the exponential growth model still generally holds for such populations, we state that the exponential growth model is fairly insensitive to changes in the age distribution. This assumption will be validated later on by comparison of predicted with observed population values.

If we take the assumption addressed in the previous paragraph to be true, then our primary assumption likely holds if changes in age distribution do not significantly impact the average mortality rate of the disease. We assume this to be true since rinderpest generally inflicts young cattle, which are of the same age group and consequently have approximately the same mortality rate (6).

Therefore, we assume that the wildebeest population growth rate can be approximated by the differential equation in 5, with P being population, $\frac{dP}{dt}$ being the rate of population growth predicted from the best fitting growth model, k being a constant proportional to the mortality rate, and SP being rinderpest seroprevalence.

$$\frac{dP}{dt} = \frac{dP_1}{dt} - kP(SP) \quad (5)$$

No matter which population growth model we choose to model $\frac{dP_1}{dt}$ with, we must evaluate the intrinsic natural growth rate r . To do so, we assume that the period after the elimination of rinderpest and before the wildebeest population begins to plateau represents a time during which population growth is characterized primarily by an exponential growth determined by r . Since rinderpest seroprevalence reaches 0 in 1963 and reaches its maximum in 1977, we crop the rinderpest population data to this time frame and fit it with an exponential curve in order to determine r , as in 3.

Next, we must determine the carrying capacity K . The carrying capacity is difficult to determine rigorously due to the irregular oscillation of the population after 1977. We assume that the midpoint of the trendline of the cropped data on [1977, 2003] is a reasonable approximation of K . This analysis is depicted in 4.

The final parameter that we must estimate is k , a constant proportional to the wildebeest mortality rate associated with rinderpest. To do so, we assume that the population was at a state of dynamic equilibrium in 1958, justifying this by noting that the rate of population growth is low at this time. So we assume $\frac{dP_1}{dt} = kP(SP)$. Additionally, the population is well below the carrying capacity in 1958, so we assume that $\frac{dP_1}{dt}$

Wildebeest Population vs Year Cropped On Exponential

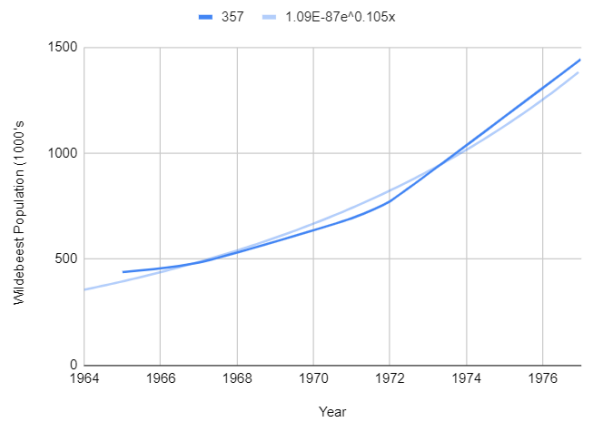


Figure 3. Wildebeest population over time from 1963 to 1977. The best fit exponential curve has an exponent of $0.105x$, suggesting that r is approximately 0.105.

Wildebeest Population vs Year Cropped Around Carrying Capacity

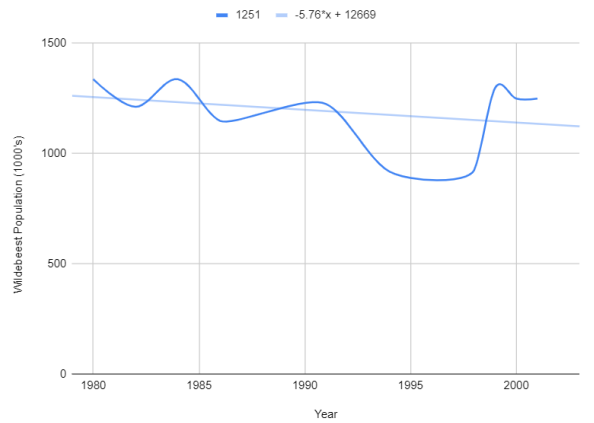


Figure 4. Wildebeest population over time from 1977 to 2003. The midpoint of the trendline represents a population of approximately 1250 thousand; the trendline is not reported with enough precision to determine this value more precisely. We therefore take K to be 1250.

is characterized more strongly by exponential growth regardless of the model in use. So we assume $\frac{dP_1}{dt} = rP$. We use the resulting equality $rP = kP(SP)$, along with our previously computed values of r and K , to estimate k to be 0.13.

At this point, we have computed all parameters necessary to extend the population growth models from 4 with the assumption from 5. We compare the resulting plots with the observed wildebeest population data to determine the best fit growth model in 5.

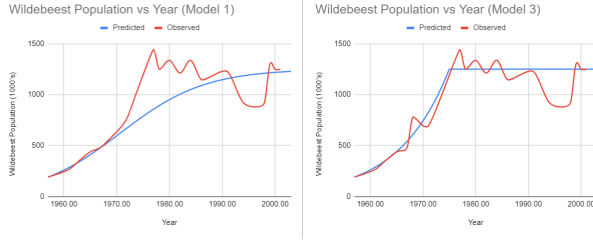


Figure 5. Comparison of Model 1 and Model 3. The plots are generated from integration of the respective differential equations using LRAM with rectangular widths of 0.01 years, and initial condition $P(1957) = 191$. Model 3 produces a clearly superior fit.

We find that Model 1 produces the superior fit, so we plug this model into 5 to get our final equation 6, with P being population (in thousands), t being time in years, and SP being rinderpest seroprevalence.

$$\frac{dP}{dt} = \begin{cases} 0.105P - 0.8P(SP) & P < 1250 \\ 0 & P \geq 1250 \end{cases} \quad (6)$$

6. WILDEBEEST POPULATION VS FIRE FREQUENCY

According to (4), the wildebeest population impacts fire frequency indirectly by changing the density of grass, which serves as a fuel for fire. So we must model the relationship between wildebeest population and grass density and between grass density and fire frequency.

We start by considering the relationship between wildebeest population and grass density. The savannah grasses of Africa grow extremely quickly. In general, “fires in the savannah burn mainly dry grasses that re-grow each year: the CO_2 released by fires in grasslands is reabsorbed by the growth of new grass the next year” (1). Since grass does not grow during the dry season (it dries and wilts, becoming fuel for fires), we assume that at some point during the wet season (the period during which wildebeest graze on the Serengeti), grass that has been burnt has the capacity to grow back to full density. We also assume that on average, wildebeest consume grasses at a constant rate proportional to the caloric intake that they require. Thus, the total amount of grass on the Serengeti is reduced from the carrying capacity by an amount proportional to the number of wildebeest. We further assume that if an individual wildebeest has the choice to graze on multiple plots of land, it will choose the one with the greatest density of grass. Aggregating the effect of this behavior, we find that the grass density on the Serengeti will be approximately uniform by the end of the grazing period (the start of the dry season).

Therefore, we claim that the relationship between grass density and wildebeest population at the end of the wet season can be summarized by 7, where G is grass density, Kg is the carrying capacity for grass density, k is a constant of proportionality, and P is the wildebeest population. We justify this equation below.

$$G = Kg - kP \quad (7)$$

Next, we attempt to determine the relationship between grass density and fire frequency. We define fire frequency in the same way that it is defined in (4) (i.e. fraction of the Serengeti burnt in a given year). We assume that the probability that a fire starts is largely independent of grass density, and depends more strongly on other factors (ex. temperature, moisture content). However, we recognize that grass density correlates directly with the amount of fuel available to an existing fire, which increases said fire’s heat and intensity. Thus, we assume that the probability that a fire spreads is directly proportional to the grass density. We further assume that a plot of land that has caught fire will continue to burn for an amount of time before the fire is extinguished. Once a plot of land has been burnt, it will not catch fire again or spread fire to other plots of land for the rest of the season (assume its remaining store of fuel is insufficient to sustain another fire). Based on these assumptions, it is intuitively clear that when there is a lot of land available for burning, the relationship between grass density and fire frequency is approximately geometric (exponential), as the probability of a fire spreading from one point to another is inherently related to the n th product of the probability that the fire spreads. However, as a greater proportion of the Serengeti is burnt, patches of burnt grass limit the spread of fire, so the rate at which fire frequency changes with an increase in grass density begins to taper off. Since fire frequency is defined as the proportion of the Serengeti that burns in a given year, it is clear that fire frequency can not exceed 1. Thus, one may predict that the relationship between grass density and fire frequency is characterized by a logistic growth model. We therefore predict that an equation relating grass density and fire frequency will be of the form in 8, where F is fire frequency, G is grass density, and r is the intrinsic rate of natural increase.

$$\frac{dF}{dG} = rF(1 - F) \quad (8)$$

Since no data for grass density on the Serengeti plain was available, 8 could not be validated by means of comparison with real-world data. To mitigate this issue, a program was developed to simulate the spread of fire

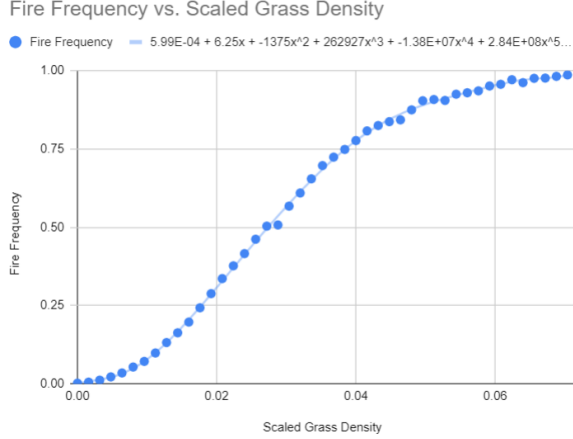


Figure 6. Synchronous cellular automaton results. The evolution of fire frequency with grass density strongly exhibits the characteristic shape of a logistic curve, supporting the model in 8.

during a dry-season on the Serengeti plain on a day-by-day basis. The spread of fire is modeled as a synchronous cellular automaton running on a grid of size 650 by 650. The length of a dry-season is five months long, which is approximately 150 days, so the simulation was run for 150 generations (2). The number of fires on the Serengeti plain ranged from 474 to 1456 between 2001 and 2014 (3). Thus, we assume that the average number of fires on the Serengeti is 965, the average of these extremes. The simulation was run for 45 arbitrarily scaled grass densities spaced out by regular intervals. The simulation outputs a fire frequency that is an average of three trials for each grass density value, since each individual run can have significant variability. These results support our mode, and are depicted in 6.

To relate wildebeest population to fire frequency, we must combine the predicted wildebeest-grass and grass-fire relationships. We start by differentiating 7 with respect to population to get the equation for the rate of change of grass density with population in 9, where G is grass density, P is wildebeest population, and k is a positive proportionality constant.

$$\frac{dG}{dP} = -k \quad (9)$$

We can now determine $\frac{dF}{dP}$ to be $\frac{dG}{dP} \times \frac{dF}{dG}$ by applying the chain rule, so we get equation 10, where F is fire frequency, P is wildebeest population, and c is a positive proportionality constant.

$$\frac{dF}{dP} = cF(1 - F) \quad (10)$$

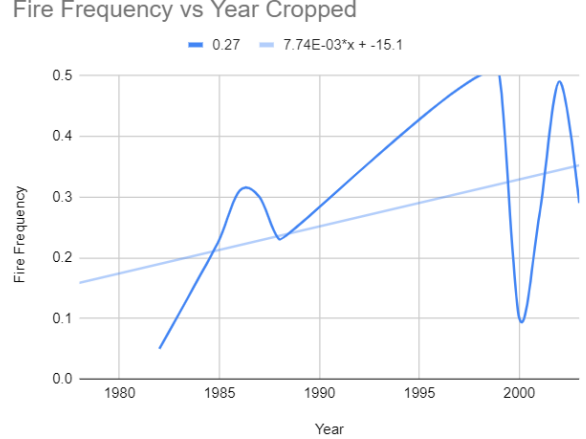


Figure 7. Fire frequency over time from 1977 to 2003. The best fit line has a midpoint of 0.25.

The value of c was determined so that when P is near the carrying capacity K , $F(P)$ approximates the average fire frequency over the time when the wildebeest population had stabilized. Since the wildebeest population is stable over the years 1977 to 2003, we estimate this average fire frequency to be the midpoint of the trend-line over this time frame, as depicted in 7.

Combined with the initial condition $F(190.8) = 0.771$, the value of c was determined to be 0.0022. To confirm that our methodology yielded reasonable results, a graph comparing our predicted relationship between wildebeest population and fire frequency with the observed relationship was generated, and is depicted in 8.

The predicted fire frequency appears to model the observed fire frequency fairly well. Note that there are many factors beyond wildebeest population that influence fire frequencies, resulting in “random” variation that can not be accounted for with wildebeest population data alone. For example, annual variation in rainfall can have a major impact on fire frequencies. It is also worth noting that the fire frequency relationship above would not continue to hold as populations get progressively larger, as cattle farmers burn large portions of the savannah as part of their savannah management technique, keeping the fire frequency above a certain minimum. Despite these limitations, the relationship between fire frequency and population can be approximated by equation 11, where F is fire frequency and P is wildebeest population in thousands.

$$\frac{dF}{dP} = -0.0022F(1 - F) \quad (11)$$

7. FIRE FREQUENCY VS TREE DENSITY

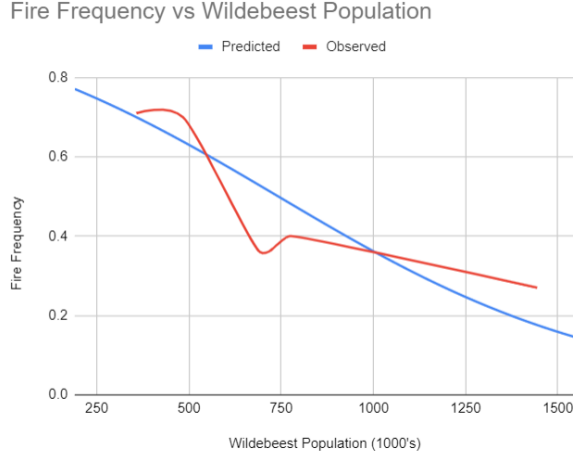


Figure 8. Predicted vs observed relationship between wildebeest population and fire frequency. The prediction plot was generated by numerical integration of 10 using LRAM with 0.05 thousands and initial condition $F(190.8) = 0.771$.

The relationship between fire frequency and tree density (ha^{-1}) is analogous to the relationship between rinderpest and wildebeest. In general, both the wildebeest and tree populations can be considered to grow according to a standard growth model, but with a counteracting factor reducing the growth rate, as in 4. However, an important distinction between the rate of growth of trees and that of wildebeest is that trees tend to conform to model 1 as opposed to model 2, since trees have extremely long lifespans and relatively low fertility rates (4.2).

In the case of the relationship between fire frequency and tree density, we assume that the amount of tree matter that burns in a given year is proportional to the product of tree density and the amount of the Serengeti that burned that year. In other words, we assume that the rate of change of tree density is reduced by an amount proportional to the fire frequency. We assume that the proportionality constant is less than 1 since trees are more fire resistant than grasses, so grasses comprise the greater proportion of the Serengeti that burns in a year. Thus, we conclude that the relationship between fire frequency and tree density can be approximated by 12, where T is tree density, t is time, r is the intrinsic rate of natural increase, K is the carrying capacity, k is a sub-unity proportionality constant, and F is fire frequency.

$$\frac{dT}{dt} = rT\left(\frac{K - T}{K}\right) - kFT \quad (12)$$

It is particularly difficult to determine the values of the three constants r , k and K due to limited data. Only one data point exists for the observed tree density, and

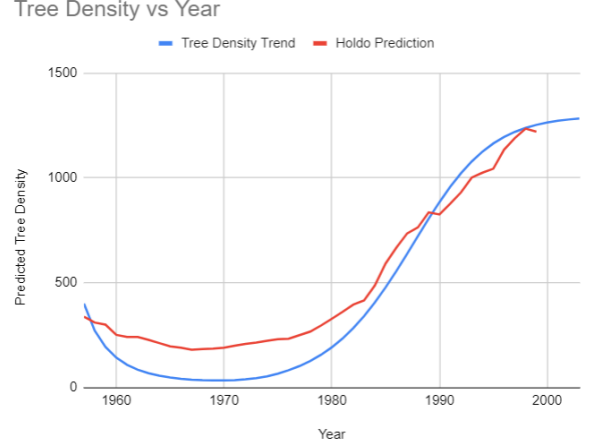


Figure 9. Sample tree density over time plot. The graph is generated by numerical integration of 12 with initial condition $T(1999) = 1260$, $r = 0.5$, $k = 0.975$, and $K = 2500$.

the predicted tree density provided by (4) at this point differs by over 200 ha^{-1} . Due to these limitations, we provide an arbitrary such graph with initial condition $T(1999) = 1260$ in 9.

We see that both graphs exhibit similar trends, which offers some support for 12 as a model for the relationship between fire frequency and tree density. However, the full extent to which 12 models the actual tree densities over time is very difficult, if not impossible, to determine given the available data.

8. CONCLUSION

Although the Serengeti is a complex ecosystem with many interdependent factors, we find that there exist strong causal relationships (rinderpest seroprevalence to wildebeest population, wildebeest population to grass density, grass density to fire frequency, fire frequency to tree density) that can be used to link each of these variables together with reasonable accuracy (with the exception of tree density, which remains ambiguous due to a lack of data). We have already aggregated the effects of most of these variables to approximate the value of each variable over time, enabling comparison with the observed results in the format in which they were originally provided. We further demonstrate the strength of the relationships determined in this solution by aggregating the expected effects of rinderpest seroprevalence and wildebeest population to yield an estimate for fire frequency over time in 10.

As with the other relationships determined in this paper, this relationship does not capture the variability in

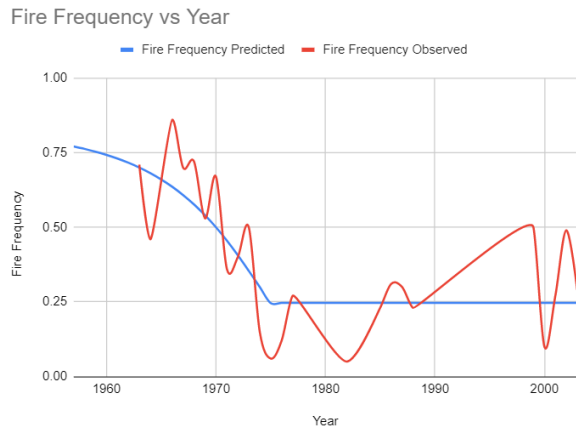


Figure 10. Predicted fire frequency over time. The plot was generated by aggregating the results predicted in 5 and 8.

the observed results, but it successfully quantifies the general trends of the observed data.

REFERENCES

- [1]Colin Beale. African Grasslands Are Meant to Burn – We Can't Let This Distract from the Amazon Fires.
- [2]Philip Briggs. Serengeti Weather and Climate.
- [3]James R. Probert et al. Anthropogenic modifications to fire regimes in the wider Serengeti-Mara ecosystem.
- [4]Ricardo M. Holdo et. al. A Disease-Mediated Trophic Cascade in the Serengeti and its Implications for Ecosystem C.
- [5]Green. ESS Topic 8.4 Human Population Carrying Capacity.
- [6]Anna R. Spickler. Rinderpest Cattle Plague.
- [7]Joel Trexler. Density Dependent Population Growth.
- [8]John Vandermeer. How Populations Grow: The Exponential and Logistic Equations.