

Modeling World Population

Dalton Deshotel

May 17, 2013

Abstract

A survey of population modeling first order differential equations and systems of differential equations. Using data available from 1900 to 2010, each model's predictive accuracy is analyzed and discussed. This survey uses MATLAB experimentation and numerical and analytic methods for solving first order differential equations.

1 Models/Methods Used

1.1 Data Sources

Population Growth and Earth's Human Carrying Capacity, by Joel E. Cohen is the source of the logistic model with variable carrying capacity. [1] Data from 1950 to 2010 comes from the United Nations Population Information Network. [2] Pre-1950 data can be found from the US Census Bureau. [3]

1.2 Exponential

For sake of completion, the exponential (Malthusian) model is included in the survey. Consider a population, $P(t)$, at any time t . The exponential model is derived by assuming that the change in population, $P'(t)$ is proportional to the population times the net growth rate (birth rate minus death rate in population).

$$b = \text{birth rate}, d = \text{death rate} \longrightarrow r = b - d$$
$$P'(t) = r \cdot P(t)$$

This first order differential equation is easily separated and solved (population is always positive):

$$\frac{dP}{dt} = r \cdot P \longrightarrow \frac{1}{P} \cdot dP = r \cdot dt$$
$$\int \frac{1}{P} \cdot dP = \int r \cdot dt$$
$$\ln(P) = r \cdot t + C \longrightarrow P = e^{r \cdot t + C} = e^C \cdot e^{r \cdot t}$$

It can easily be shown that $e^C = P_0$, the initial population. This results in the complete form of the familiar exponential population model:

$$P(t) = P_0 \cdot e^{r \cdot t}$$

The most obvious flaw using this method to model population growth is that it is unbound. For modeling populations of living organism, this means that the function will eventually exceed possible numbers.

1.3 Logistic

1.3.1 Verhulst Model

First published by Pierre-François Verhulst, the logistic model attempts to model real-world population dynamics by adding a **carrying capacity**, denoted by K , which provides a theoretical maximum limit value for the population.

The logistic model is derived by modifying growth and death rates to vary in proportion to the size of the population. A relatively large population has to fight over scarcer resources, while a small population can grow more rapidly. Generally, the greater the population the lower the birth rate, and greater the death rate. By adding in variable birth and death rates, the logistic population model is derived:

$$\begin{aligned} d + a \cdot P &= \text{death rate}, \quad b - c \cdot P = \text{birth rate} \\ P'(t) &= [(b - c \cdot P) - (d + a \cdot P)] \cdot P = [b - d - (a + c) \cdot P] \cdot P \end{aligned}$$

This is simplified by letting the **natural reproductive rate**, $r_0 = b - d$, and then $a + c = r_0/K$ where $K = r_0/(a + c)$. These get plugged into the equation:

$$P' = \left(r_0 - \frac{r_0 \cdot P}{K} \right) \cdot P = r_0 \cdot \left(1 - \frac{P}{K} \right) \cdot P$$

Since $P'(t) = 0$ when $P = K$, K is called the **carrying capacity** of the population. In population biology, K and r_0 form the basis of r/K selection theory.

This logistic equation is autonomous, and separable. It can be written as:

$$\begin{aligned} \frac{K \cdot dP}{P \cdot (K - P)} &= r \cdot dt = \frac{1}{P} + \frac{1}{K - P} \\ \left(\frac{1}{P} + \frac{1}{K - P} \right) \cdot dP &= r \cdot dt \end{aligned}$$

After integrating both sides:

$$\begin{aligned} \ln||P|| - \ln||K - P|| &= r \cdot t + C \\ \longrightarrow P(t) &= \frac{K \cdot A \cdot e^{r \cdot t}}{1 + A \cdot e^{r \cdot t}} \end{aligned}$$

1.3.2 Augmented Logistic

In his 1995 paper *Population Growth and Earth's Human Carrying Capacity*, Joel E. Cohen uses an augmented logistic with variable carrying capacity. [1]

This is the version of the logistic equation used by Cohen:

$$P' = r \cdot P(t) \cdot [K(t) - P(t)]$$

The second equation in this system models carrying capacity as inversely proportional to population and proportional to the rate of change of population:

$$K' = \frac{L}{P(t)} \cdot P'$$

The second equation in this system models the capacity of each additional person to alter the carrying capacity of the Earth. The smaller the population, the more each individual can affect the carrying capacity. However, when the population is very large, the equation $K' = \frac{L}{P(t)} \cdot P'$ nears $K' = 0$, which reduces to the special case where K is constant, in which case the standard logistic model can make a good approximation.

1.4 Using r/K theory to think about models

Although r/K theory is a theory of evolutionary adaptation, and has since become only one aspect of other, more complex theories, the core concept of the theory—the interplay between growth and carry capacity—can serve as a guide when thinking about human populations. Essentially, the theory states that there are two general reproductive strategies that a species implements to variable degrees:

- First, there's the **r-strategy**, when the species breeds in large numbers but consequently has fewer resources to ensure the survival of each individual offspring.
- Second, there's the **K-strategy**, when a species breeds in smaller numbers, and therefore has more resources to devote to each offspring.

It is important to note that r and K strategies are not mutually exclusive, and that all species can be thought of as having varying degrees of r and K-selected traits. For thinking about human population growth, this means accounting for how much an individual alters the carrying capacity—which is the core concept behind Cohen's augmented model. [1]

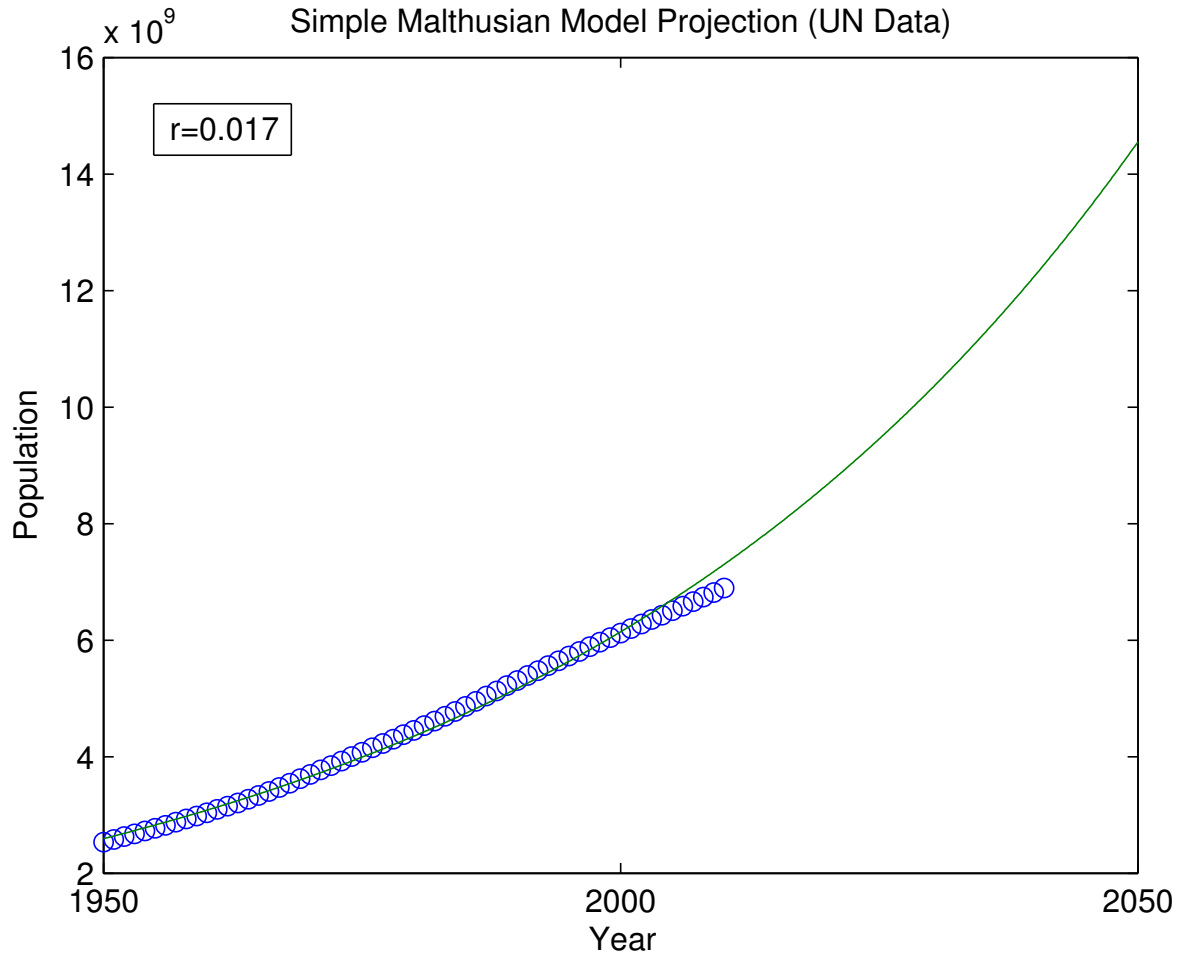
1.5 Polyfit

Using the available data, the Matlab function polyfit will provide a point of comparison for the various differential equations surveyed. This illustrates how effective these methods are compared to basic curve fitting.

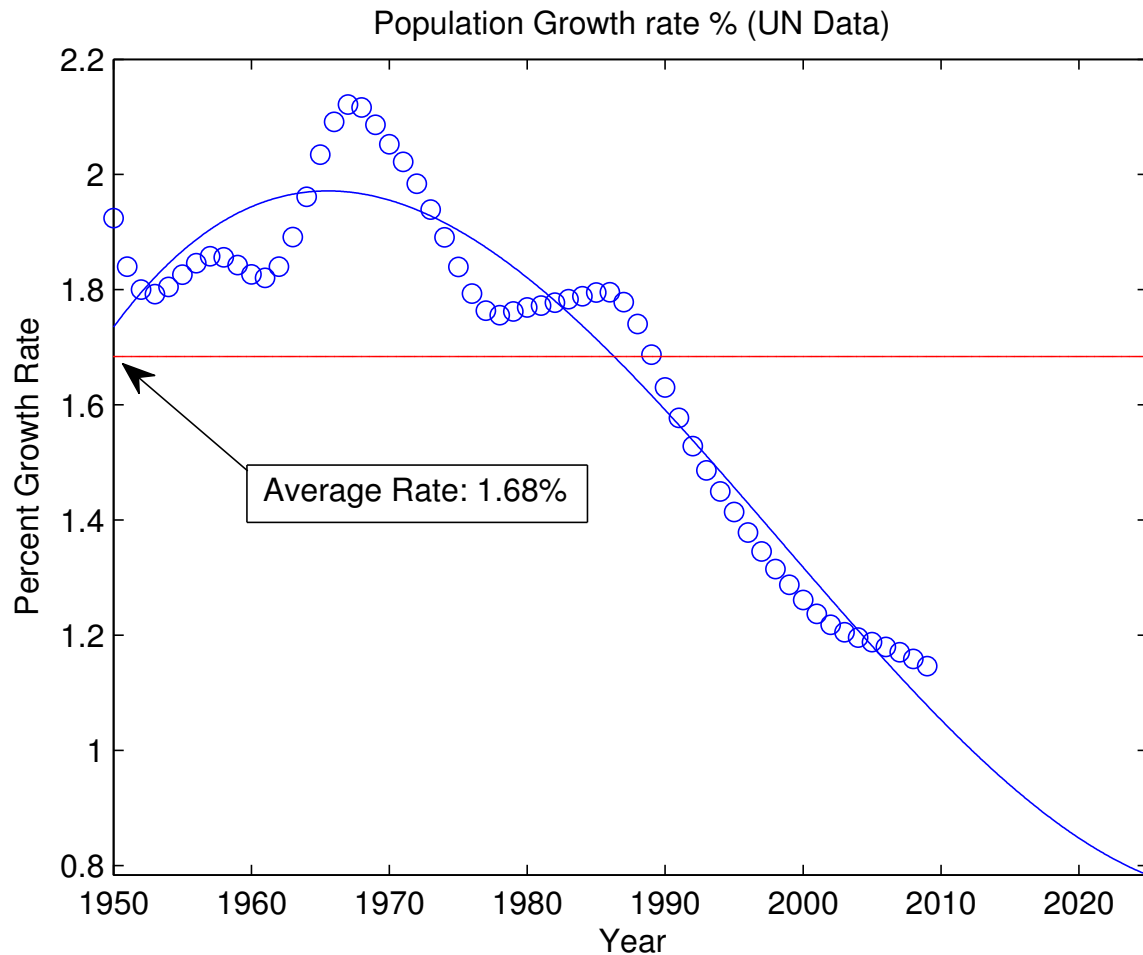
2 Experiments

2.1 Exponential

While the basic exponential Malthusian model can predict short-term exponential growth, it fares much more worse for long-term projections. The exponential function was derived using the log of the dataset, and line-fitting using a first degree polynomial and the polyfit function in MATLAB:



Looking at the growth rates between 1950 and 2010, it shows a variable growth rate that isn't factored in by the exponential function alone. Population growth rate peaked in the mid 1960's, and has, generally, been decreasing ever since:

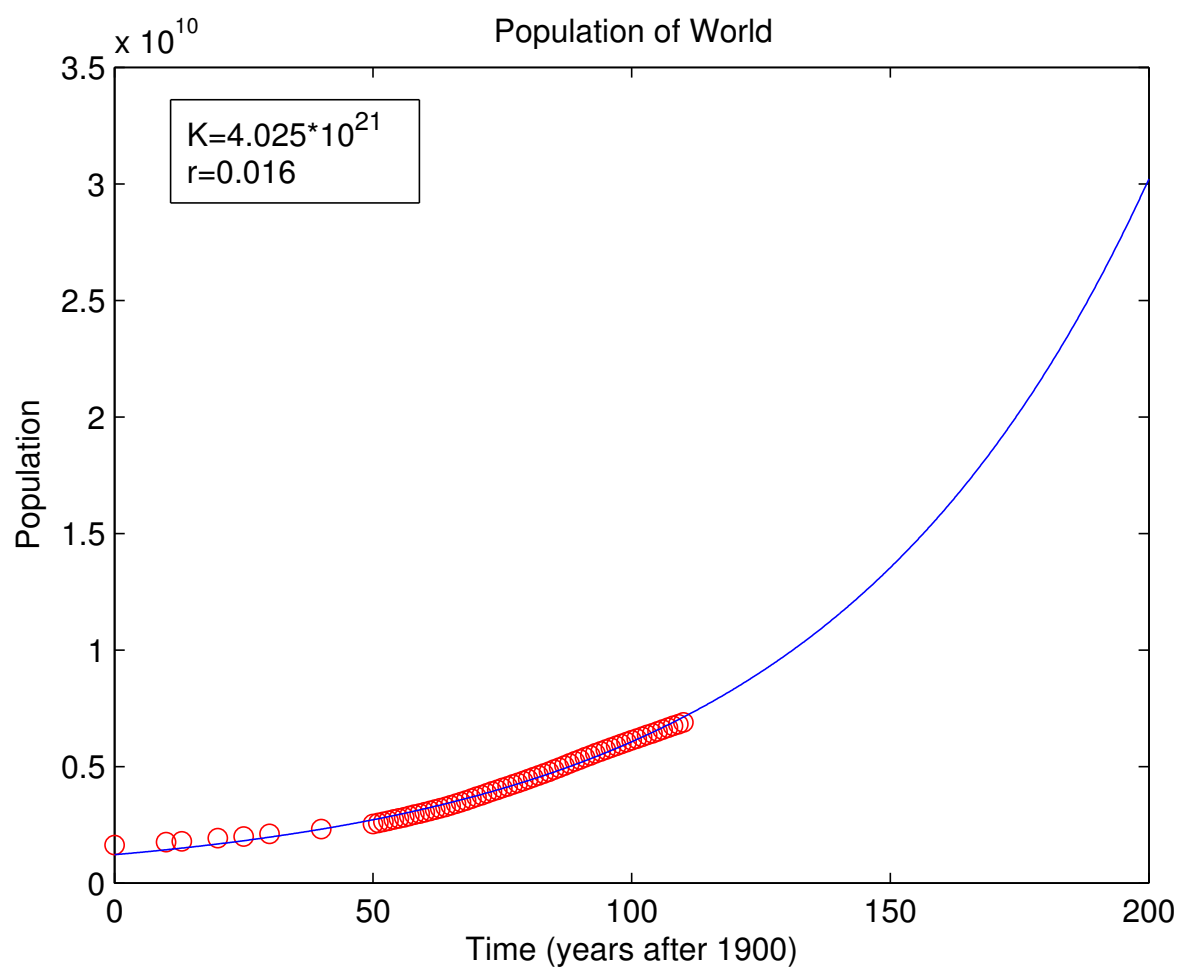


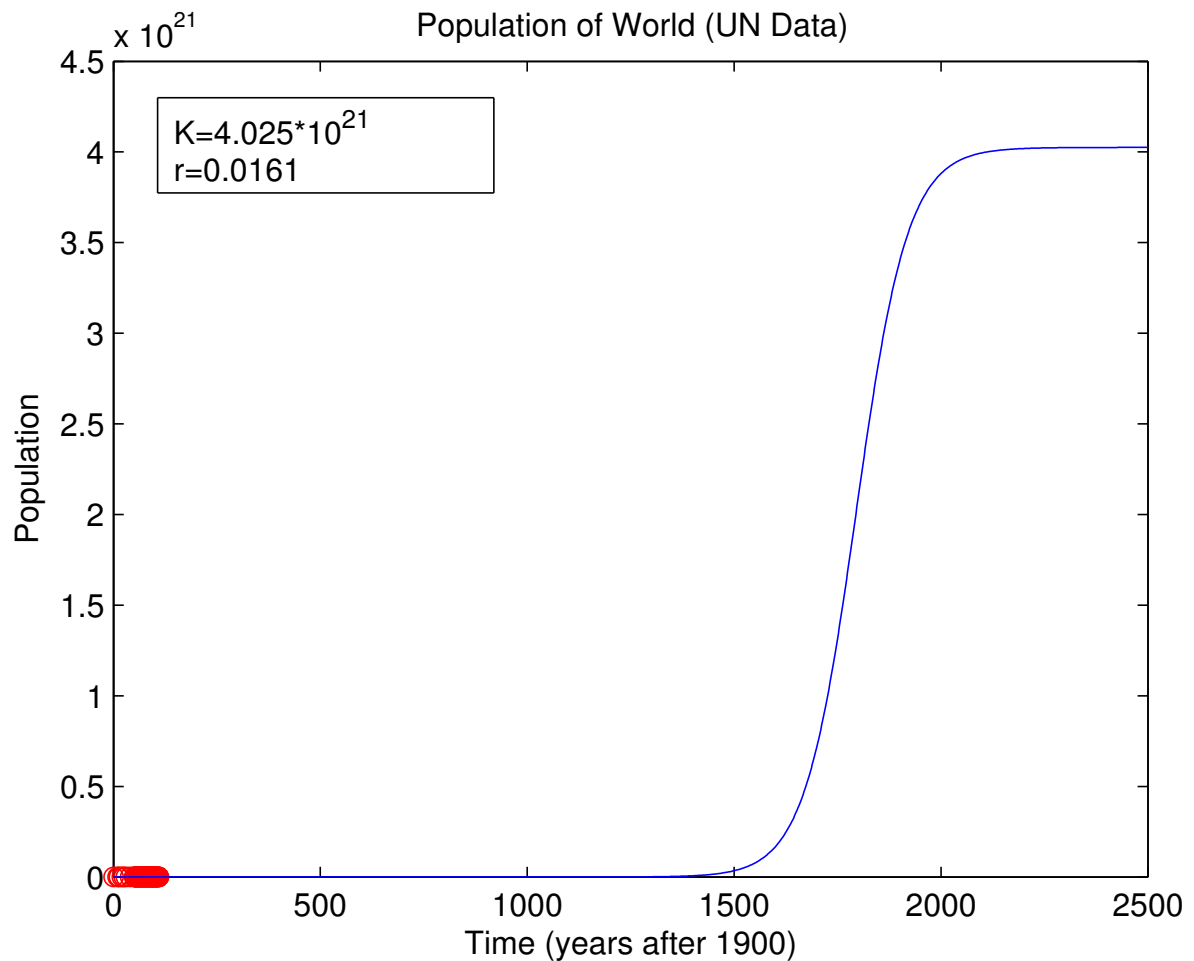
The data shows both an increasing and decreasing rate of growth—an inflection point. The Malthusian model—which models only static growth rate—is not a very good fit for the population data. A far better model would be one that factors in this changing rate of growth.

2.2 Logistic

2.2.1 Logistic Modeling Using Data from 1900

At first glance, this models seems to have given a reasonable projection for the population in the year 2050: $P \approx 15 \cdot 10^9$. However, using this data the model produces the unrealistic K-value of $4.025 \cdot 10^{21}$, which would take approximately 2,000 years to reach. For comparison, the diameter of Milky Way Galaxy is approximately 10^{21} meters wide!



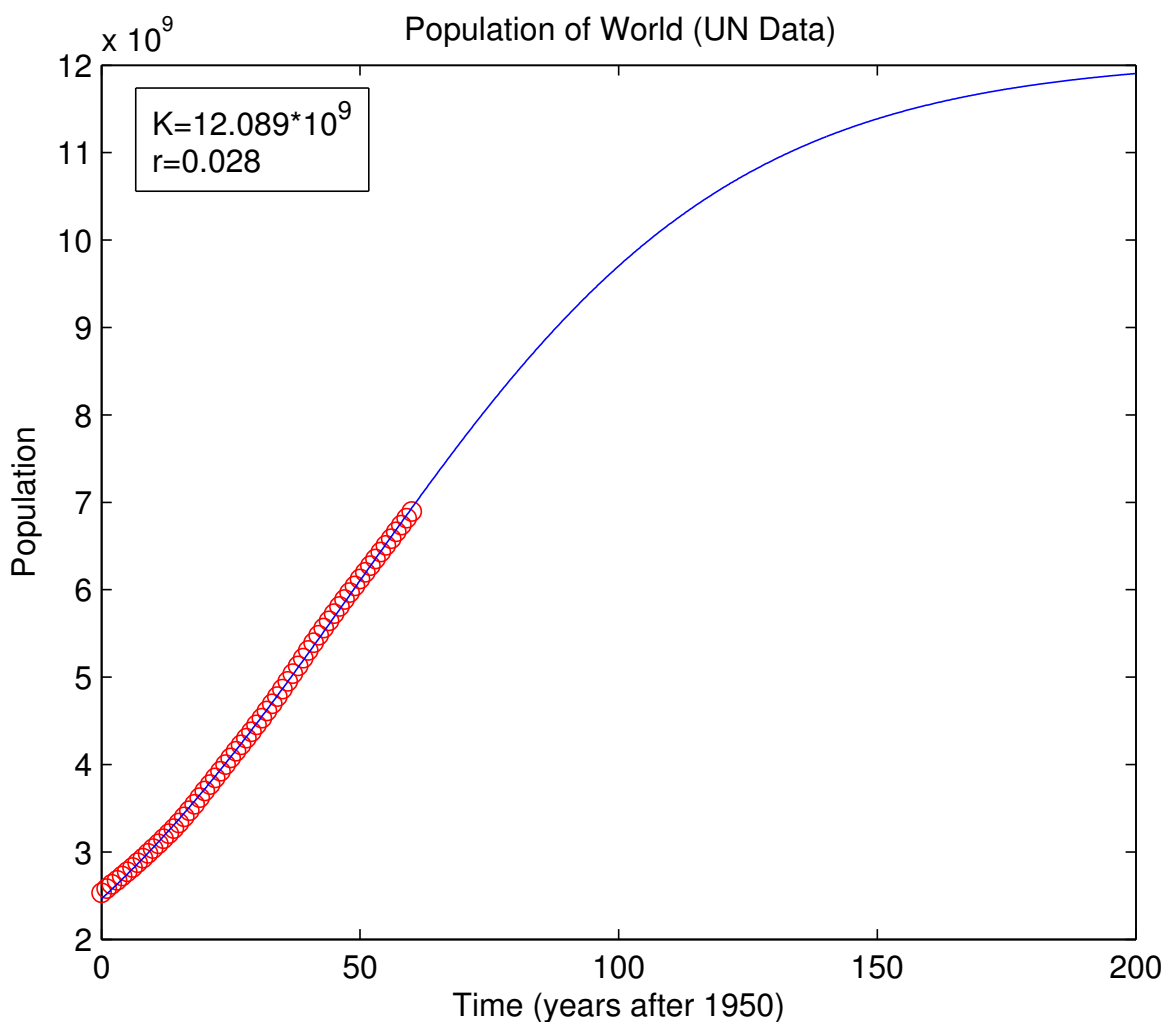


This example clearly shows the limitations of the logistic function when modeling human population. Unlike bacteria in a petri dish, or fish in a lake, humans have an evolving relationship with the environment. In order for this model to work, many of the same assumptions that are made for animal populations are necessary:

- A population's ability to use resources is relatively stable in the short term. Animals do not alter their environment through technological means such as agriculture, civil engineering, etc.
- When a population is near its carrying capacity, each individual consumes just enough resources to survive and reproduce. That is, considerations such as quality of life aren't factored in. Maximum carrying capacity should be thought of as strictly a theoretical maximum, not a predicted future population.
- Distribution of resources in a small environment can be more uniform than that of a large one. One example of this is the discrepancy between first and third world poverty levels. Therefore, a simple mathematical model must assume optimum resource distribution.

2.2.2 Logistic Modeling Using Data from 1950

Plugging in only the numbers from 1950, the model produces the much saner theoretical carrying capacity of $12 \cdot 10^9$ —which agrees with other estimates (Cohen). [?] As the growth rate data shows, this time period has both increasing, then decreasing rates of change. Using data from this period models a world that is just beginning to encounter the problems of rapid population growth: pollution, crowding, etc. Including pre-1950 data factors in a growth period when humans weren't as technologically capable of a truly globe-spanning civilization as they were in the post-1950 era.



Considering the augmented logistic system, and the function $K' = \frac{L}{P(t)} \cdot P'$, it's reasonable to assume that as P increases, the capacity of each individual to alter the carrying capacity diminishes to almost nothing. And, as human population increases, the basic logistic function, often thought to be ill-suited for modeling human population, may actually increase in accuracy with each passing year.

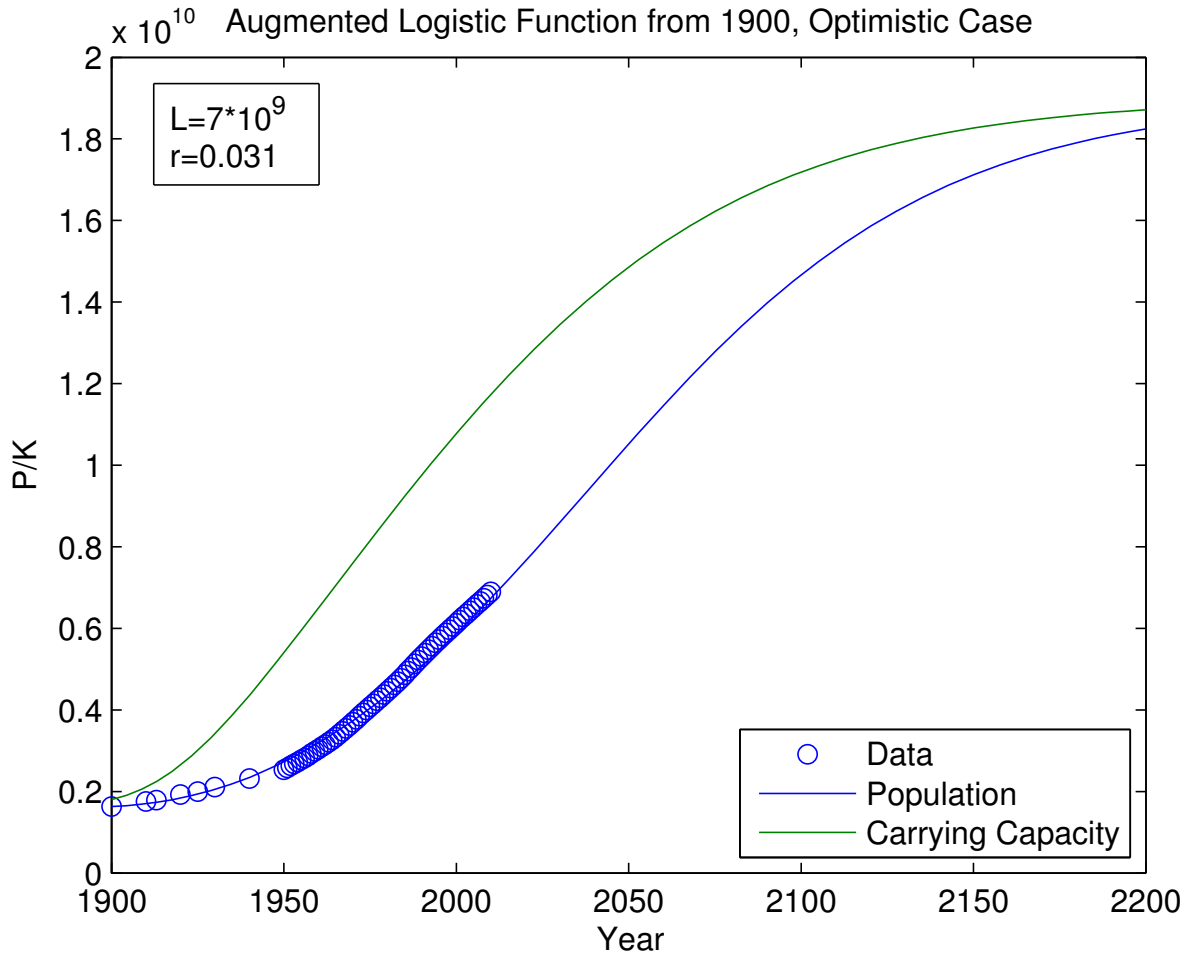
2.2.3 Augmented Logistic Modeling

The following graphs were made using the ode45 function in MATLAB, and manual tweaking of the r/K numbers derived from previous experiments, and they use the following system of first order differential equations derived from Cohen's 1995 paper:

$$P' = r \cdot P \cdot (1 - P/K),$$

$$K' = \frac{L}{P} \cdot P'$$

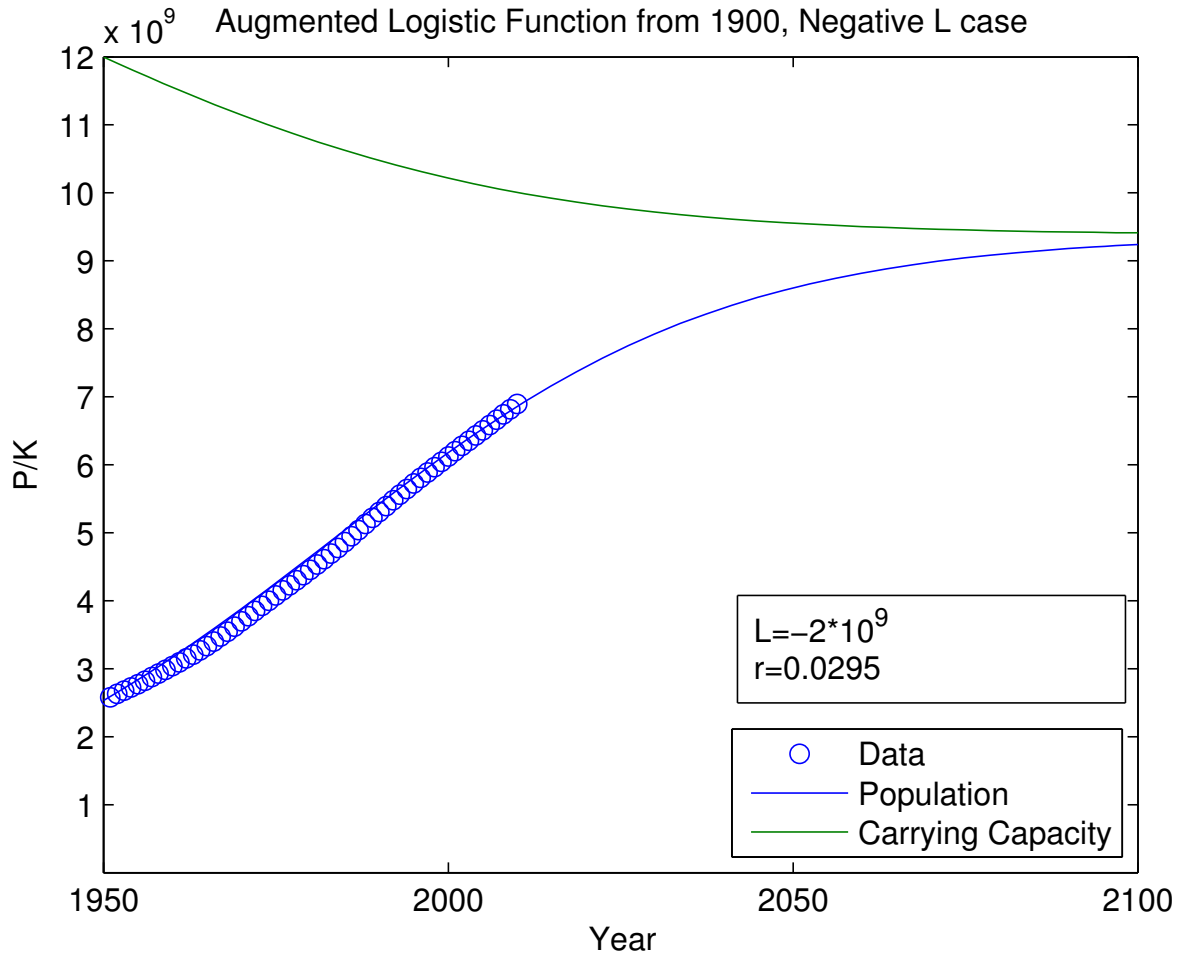
The first experiment used the data from 1900. Using this system the carrying capacity maxes out at $K \approx 1.9 \cdot 10^9$. Using a positive L-number makes the assumption that each additional human being, on average, can increase the carrying capacity, albeit at an ever-decreasing rate:



This model provides a much saner fit for the data than the basic logistic.

The second experiment used the data from 1950, and the initial carrying capacity of $12 \cdot 10^9$.

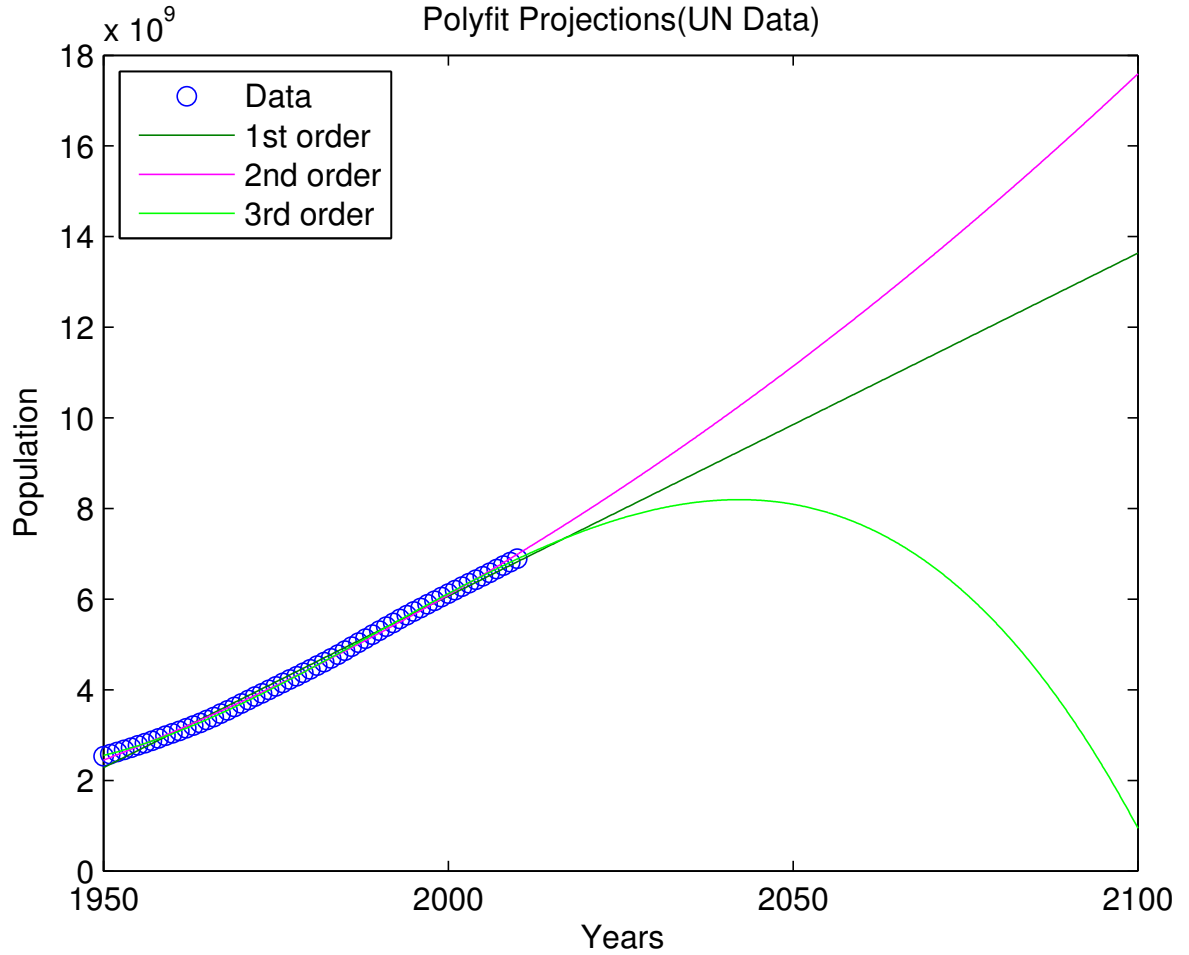
The negative L-number assumes that each additional person negatively alters the carrying capacity:



This produces an eventual carrying capacity of $K \approx 9.5 \cdot 10^9$. Considering many of the issues of pollution, overcrowding, poverty and so forth, this is a more sobering, and realistic scenario. With $12 \cdot 10^9$ as the theoretical maximum from the standard logistic model, it is unlikely that human population can exceed this number in the foreseeable future.

2.3 Polyfit

For comparison, the MATLAB function polyfit:



Each of these polynomials predict unrealistic future populations. Even the simple Malthusian model does a better job of fitting the data. At best, a first degree polynomial can give a decent short-term approximation of population growth.

3 Discussion/Conclusions

These methods are just a few of the many mathematical models available for predicting future human population growth. However, they can only provide frames of reference for future planners for what to expect if current trends do not change.

A large shortage of a once-abundant resource such as petroleum, habitable land, etc. can easily thwart any of these models. For a more serious study of the population dynamics of the Earth, these models need to be used in conjunction with real-world data from various sources. There is also the question as to whether or not a species capable of modeling its own population dynamics can alter its reproductive habits to avoid a Malthusian-style population crash. That question, however, is beyond the scope of this paper.

References

- [1] Cohen, J. E. (1995, July 21). Population Growth and Earth's Human Carrying Capacity. *Science*, 269(5222).
- [2] United Nations Population Information Network. Retrieved 2013, from <http://www.un.org/popin/>
- [3] World Population. Retrieved May 16, 2013, from http://www.census.gov/population/international/data/worldpop/table_population.php

4 MATLAB code

4.0.1 Malthusian Model

```
Q=[1950 1951 1952 1953 1954 1955 1956 1957 1958 1959 1960 1961 1962 1963 1964 1965
1966 1967 1968 1969 1970 1971 1972 1973 1974 1975 1976 1977 1978 1979 1980 1981 1982
1983 1984 1985 1986 1987 1988 1989 1990 1991 1992 1993 1994 1995 1996 1997 1998 1999
2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010];
M=(103)*[25322292580960262844826757662723726277288228235132875642292906929834353038413309
T=Q-linspace(1950,1950,length(Q))
G=log(M)
L=polyfit(T,G,1)
a=L(1)
b=L(2)
t=linspace(1950,2050,1000);
P=exp(b)*exp(a*(t-1950))
plot(Q,M,'o',t,P)
xlabel('Year')
ylabel('Population')
title('Simple Malthusian Model Projection (UN Data)')
```

4.0.2 Growth Rates

```
close all
P=[2532229 2580960 2628448 2675766 2723726 2772882 2823513 2875642 2929069 2983435
3038413 3093909 3150242 3208212 3268896 3333007 3400823 3471955 3545613 3620652 3696186
3772048 3848319 3924668 4000764 4076419 4151410 4225864 4300402 4375899 4453007 4531799
4612120 4694097 4777828 4863290 4950591 5039478 5129113 5218375 5306425 5392939 5478009
5561744 5644416 5726239 5807212 5887260 5966465 6044931 6122770 6200003 6276722 6353196
6429758 6506649 6583959 6661637 6739610 6817737 6895889]
t=linspace(1950,2025,1000)
R=[];
for i=2:length(P)
R=[R abs((P(i)-P(i-1))/P(i-1))];
end
T=1950:(1950+length(R)-1);
plot(T,100*R,'o');
hold on
p=polyfit(T,R,3);
f=polyval(p,t);
plot(t,100*f,t,100*0.01684,'r');
axis([1950 2025 min(100*f) 2.2]);
title('Population Growth rate
hold off
xlabel('Year')
ylabel('Percent Growth Rate')
```

4.0.3 Logistic

figure

close all

```
Q=[1900 1910 1913 1920 1925 1930 1940 1950 1951 1952 1953 1954 1955 1956 1957 1958  
1959 1960 1961 1962 1963 1964 1965 1966 1967 1968 1969 1970 1971 1972 1973 1974 1975  
1976 1977 1978 1979 1980 1981 1982 1983 1984 1985 1986 1987 1988 1989 1990 1991 1992  
1993 1994 1995 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009  
2010]
```

```
M=(103)*[16288301753594179109119280172000000210918223197712532229258096026284482675766272
```

```
T=Q-(linspace(1900,1900,length(Q)))
```

```
T=T';
```

```
M=M';
```

```
r=linspace(0.001,1,210);
```

```
t0=linspace(0,210,210);
```

```
[r,t0]=meshgrid(r,t0);
```

```
[m,n]=size(r);
```

```
e=zeros(size(r));
```

```
for i=1:m
```

```
for j=1:n
```

```
e(i,j)=myerror([r(i,j);t0(i,j)],T,M);
```

```
end
```

```
end
```

```
min=fminsearch(@myerror,[0.1;200],[],T,M);
```

```
r=min(1);
```

```
t0=min(2);
```

```
H=1./(1+exp(-r*(T-t0)));
```

```
K=(H'*M)/(H'*H);
```

```
figure
```

```
t=linspace(0,200,1000);
```

```
y=K./(1+exp(-r*(t-t0)));
```

```
plot(T,M,'ro',t,y)
```

```
title('Population of World (UN Data)')
```

```
xlabel('Time (years after 1900)')
```

```
ylabel('Population')
```

4.0.4 Augmented Logistic Function

```
function Fv=cohen(t,Y);
```

```
L=-2*109
```

```
r=0.0295
```

```
Fv(1,1)=r*Y(1)*(1-(Y(1)/Y(2)))
```

```
Fv(2,1)=(L/(Y(1)))*(r*Y(1)*(1-(Y(1)/Y(2))))
```

4.0.5 Augmented Logistic Plot

close all

```
Q=[1900 1910 1913 1920 1925 1930 1940 1950 1951 1952 1953 1954 1955 1956 1957 1958  
1959 1960 1961 1962 1963 1964 1965 1966 1967 1968 1969 1970 1971 1972 1973 1974 1975  
1976 1977 1978 1979 1980 1981 1982 1983 1984 1985 1986 1987 1988 1989 1990 1991 1992  
1993 1994 1995 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009  
2010]
```

```
M=(103)*[16288301753594179109119280172000000210918223197712532229258096026284482675766272
```

```
plot(Q,M,'o')
```

```
hold on
```

```
[t y]=ode45(@cohen,[1950 2100],[2.5322*109 12 * 109])
```

```
plot(t,y(:,1),t,y(:,2))
```

```
legend('Data','Population','Carrying Capacity',4)
```

```
xlabel('Year')
```

```
ylabel('P/K')
```

```
title('Augmented Logistic Function from 1900, Negative L case')
```

```
axis([1950 2100 2*109 12*109])
```