So what is a mathematical model?

Biologists typically build a descriptive model of a phenomena to hypothesize how it works, or try to estimate how an unforeseeable event could affect what they are studying. To obtain a quantifiable description of the system under study, they take their descriptive model a step further and create a “mathematical model.” A mathematical model is a mathematical representation of a relationship between “objects” in the real world. It typically describes a system by a set of variables and a set of equations that establish relationships between the variables. The variables represent some properties of the system, for example, measured system outputs often in the form of pH levels, Absorption amounts, size, time distance, etc. The actual model is the set of functions that describe the relations between the different variables. Models are used throughout the sciences such as in biology, chemistry, engineering and physics, as well as in other fields such as economics, sociology and political science.

The essence of a good mathematical model is that it is simple in design and exhibits the basic properties of the real “system” that we are attempting to understand. The model should be testable against empirical data and comparisons of the model to the real world should ideally lead to improved mathematical models. The model may suggest improved experiments to highlight a particular aspect of the problem under investigation, which in turn may improve the collection of data.

During the process of building a mathematical model, the modeler will decide what factors are relevant to the problem and what factors can be de-emphasized. Once a model has been developed and used to answer questions, it should be critically examined and often modified to obtain a more accurate reflection of the observed reality of that phenomenon. In this way, mathematical modeling is an evolving process; as new insight is gained, the process begins again as additional factors are considered. Generally the success of a model depends on how easily it can be used and how accurate are its predictions.

Some Examples of Mathematical Models in Biology:

1. **pH Model:**
   \[ \text{pH} = - \log[H^+] \]

2. **Half-Life Model:**
   \[ t_{1/2} = \frac{\ln(2)}{r} \approx \frac{0.693}{r} \]

3. **Hardy-Weinberg Law (Model):**
   \[ (p+q)^2 = p^2 + 2pq + q^2 = 1, \text{ and } p+q = 1 \]
4. **Michaelis-Menton Model for Enzyme Kinetics:** \[ V = \frac{V_{\text{Max}} [S]}{K_m + [S]} \]

5. **MacArthur-Wilson Species-Area Law (Model):** \[ S = cA^x \]

6. **General Scaling Law (Model):** \[ Y = Y_0 M^b \]

7. **Haldane Function (Genetic Mapping Model):** \[ C_{AB} = \frac{1}{2} \left( 1 - e^{-2X_{AB}} \right) \]

8. **Population Growth Model:** \[ P = P_0 e^{rt} \]

**Building a Mathematical Model**

Building a mathematical model can be a challenging, yet rewarding task. A thorough understanding of the underlying scientific concepts is necessary for developing a complete model, however, creating a model can also lead to greater understanding of the phenomenon as well.

Although problems may require very different methods of solution, the following steps outline two general approaches to the mathematical modeling process:

**APPROACH 1 (Start with the fundamental underlying principles):**

1. **Identify the problem**, define the terms, and draw diagrams where appropriate

2. **Begin with a simple model**, stating the assumptions that you make as you focus on particular aspects of the phenomenon

3. **Identify important variables and constants** and determine how they relate to each other

4. **Develop the equation(s)** that express the relationships between the variables and constants

5. **Verify and Refine the Model:** Once the model has been developed and applied to the problem, your resulting model solution must be analyzed and interpreted with respect to the problem. The interpretations and conclusions should be checked for accuracy by answering the following questions:
   - Is the information produced reasonable?
   - Are the assumptions made while developing the model reasonable?
   - Are there any factors that were not considered that could affect the outcome?
   - How do the results compare with real data, if available?

In answering the above questions, you may need to modify your model. This refining process should continue until you obtain a model that agrees as closely as possible with the real world observations that you have set out to model.
APPROACH 2 (Start with measured data):

1. Identify important variables and parameters (constants) (Dependent and independent?)
2. Obtain empirical data
3. Fit the data to an appropriate function (obtain the approximate algebraic fit to the data)
4. Develop the Model to fully express the relationships between the variables and constants.
5. Verify and Refine the Model: Same as APPROACH 1 above

Variables and Parameters

Mathematical models typically contain three distinct types of quantities: output variables (range), input variables (domain), and parameters (constants). Output variables give the model solution. The choice of what to specify as input variables and what to specify as parameters is somewhat arbitrary and often model dependent. Input variables characterize a single physical problem while parameters determine the context or setting of the physical problem. For example, in modeling the decay of a single radioactive material, the initial amount of material and the time interval allowed for decay could be input variables, while the decay constant for the material could be a parameter. The output variable for this model is the amount of material remaining after the specified time interval.

Classifying mathematical models

Mathematical models can be classified as:

1. Linear or Nonlinear: If all the mathematical functions and constraints in the model are represented entirely by linear functions, then the model is known as a linear model. If one or more of the functions or constraints are represented with a nonlinear expression, then the model is known as a nonlinear model.

2. Deterministic or Probabilistic (stochastic): A deterministic model performs the same way for a given set of initial conditions, while in a stochastic model, randomness is present, even when given an identical set of initial conditions.

   An example of a probabilistic model would be a model predicting the genetic makeup of an offspring. All of the models listed above are examples of deterministic models. There is no uncertainty or variability built into the models. Everything is determined.

3. Static or Dynamic: A static model does not account for the element of time, while a dynamic model does.

Dynamic Models can be further classified as continuous in time vs. discrete in time

Mathematical models of time dependent processes can be split into two categories depending on how the time variable is to be treated. A continuous-in-time mathematical model is based on a set of equations that are valid for any value of the time variable. A discrete-in-time mathematical model is designed to provide information about the state of the physical system only at a selected set of distinct times.
The solution of a continuous-in-time mathematical model provides information about the physical phenomenon at every time value. The solution of a discrete-in-time mathematical model provides information about the physical system at only a finite number of time values. Continuous-in-time models have two advantages over discrete-in-time models: (1) they provide information at all times and (2) they more clearly show the qualitative effects that can be expected when a parameter or an input variable is changed. On the other hand, discrete in time models have two advantages over continuous in time models: (1) they are less demanding with respect to skill level in algebra, trigonometry, calculus, differential equations, etc. and (2) they are better suited for implementation on a computer.

A-priori information

Mathematical modeling problems are often classified as either black-box or white-box models, according to how much a-priori information is available on the system. A black-box model is where there is no a-priori information available. A white-box model (also called glass box or clear box) is a system where all necessary information is available. Practically all systems are somewhere between the black-box and white-box models, so this concept only works as an intuitive guide for approach.

Usually it is preferable to use as much a-priori information as possible to make the model more accurate. Therefore the white-box models are usually considered easier, because if you have used the information correctly, then the model will behave correctly. Often the a-priori information comes in forms of knowing the type of functions relating different variables. For example, if we make a model of how a medicine works in a human system, we know that usually the amount of medicine in the blood is an exponentially decaying function. But we are still left with several unknown parameters; how rapidly does the medicine amount decay, and what is the initial amount of medicine in blood? This example is therefore not a completely white-box model. These parameters have to be estimated through some means before one can use the model.

In black-box models one tries to estimate both the functional form of relations between variables and the numerical parameters in those functions. Using a-priori information we could end up, for example, with a set of functions that probably could describe the system adequately. If there is no a-priori information, we would try to use functions as general as possible to cover all different models.

Complexity

Another basic issue is the complexity of a model. If we were, for example, modeling the flight of an airplane, we could embed each mechanical part of the airplane into our model and would thus acquire an almost white-box model of the system. However, the computational cost of adding such a huge amount of detail would effectively inhibit the usage of such a model. Additionally, the uncertainty would increase due to an overly complex system, because each separate part induces some amount of variance into the model. It is therefore usually appropriate to make some approximations to reduce the model to a sensible size. Scientists often can accept some approximations in order to get a more robust and simple model. For example Newton's classical mechanics is an approximated model of the real world. Still, Newton's model is quite sufficient for most ordinary-life situations, that is, as long as particle speeds are well below the speed of light, and we study macro-particles only.

Model evaluation/validation

An important part of the modeling process is the evaluation of an acquired model. How do we know if a mathematical model describes the system well? This is not an easy question to answer. Usually the scientist has a set of measurements from the system, which are used in creating the model. Then, if the
model was built well, the model will adequately show the relations between system variables for the measurements at hand. The question then becomes: How do we know that the measurement data is a representative set of possible values? Does the model describe well the properties of the system between the measurement data (interpolation)? Does the model describe well events outside the measurement data (extrapolation)?

A common approach is to split the measured data into two parts: training data and verification data. The training data is used to train the model, that is, to estimate the model parameters (see above). The verification data is used to evaluate model performance. Assuming that the training data and verification data are not the same, we can assume that if the model describes the verification data well, then the model describes the real system well.

However, this still leaves the extrapolation question open. How well does this model describe events outside the measured data? Consider again Newtonian classical mechanics-model. Newton made his measurements without advanced equipment, so he could not measure properties of particles traveling at speeds close to the speed of light. Likewise, he did not measure the movements of molecules and other small particles, but macro particles only. It is then not surprising that his model does not extrapolate well into these domains, even though his model is quite sufficient for ordinary life physics.
EXAMPLE OF APPROACH 1

"In October 1838, that is, fifteen months after I had begun my systematic inquiry, I happened to read for amusement Malthus on Population, and being well prepared to appreciate the struggle for existence which everywhere goes on from long-continued observation of the habits of animals and plants, it at once struck me that under these circumstances favourable variations would tend to be preserved, and unfavourable ones to be destroyed. The results of this would be the formation of a new species. Here, then I had at last got a theory by which to work".

Charles Darwin, from his autobiography. (1876)

This often quoted passage reflects the significance Darwin affords Malthus in formulating his theory of Natural Selection. What "struck" Darwin in Essay on the Principle of Population (1798) was Malthus's observation that in nature plants and animals produce far more offspring than can survive, and that Man too is capable of overproducing if left unchecked. Malthus concluded that unless family size was regulated, man's misery of famine would become globally epidemic and eventually consume Man. Malthus' view that poverty and famine were natural outcomes of population growth and food supply was not popular among social reformers who believed that with proper social structures, all ills of man could be eradicated.

Although Malthus thought famine and poverty natural outcomes, the ultimate reason for those outcomes was divine institution. He believed that such natural outcomes were God's way of preventing man from being lazy. Both Darwin and Wallace independently arrived at similar theories of Natural Selection after reading Malthus. Unlike Malthus, they framed his principle in purely natural terms both in outcome and in ultimate reason. By so doing, they extended Malthus' logic further than Malthus himself could ever take it. They realized that producing more offspring than can survive establishes a competitive environment among siblings, and that the variation among siblings would produce some individuals with a slightly greater chance of survival.

Malthus was a political economist who was concerned about, what he saw as, the decline of living conditions in nineteenth century England. He blamed this decline on three elements: The overproduction of young; the inability of resources to keep up with the rising human population; and the irresponsibility of the lower classes. To combat this, Malthus suggested the family size of the lower class ought to be regulated such that poor families do not produce more children than they can support.

In 1761, Robert Wallace published Various Prospects of Mankind in which he argued that progress would eventually undo itself by overstocking the world with people. Thirty-seven years later, Thomas Malthus published his now famous Essay on the Principle of Population. In it, he predicted that population growth would eventually outrun food supply. This prediction was based on the idea that population, if unchecked, increases at a geometric rate, whereas the food supply could only grow at an arithmetic rate. Mathematically, any increasing geometric sequence (e.g. 1, 3, 9, 27, 81) will eventually overtake all arithmetic sequences (e.g. 10, 20, 30, 40, 50). The resulting decrease in food per person will eventually lead to subsistence-level conditions. According to Malthus, the catastrophe can only be prevented by self-restraint or vice—which for him included contraception and abortion. The geometric increase in population means that every population has a constant doubling time.

Malthus did not give a time frame for his catastrophe. Thus far, population growth has been essentially geometric as Malthus predicted. The Malthusian catastrophe, however, has not occurred, principally
because food supply growth has also been roughly geometric, not arithmetic. Furthermore, the widespread use of contraception and abortion (Malthusian vices) have, as Malthus said they could, restrained population growth significantly. In fact, current, food supply per person is several times higher than when Malthus wrote his essay.

The complete essay can be found at: http://www.ac.wwu.edu/~stephan/malthus/malthus.0.html

With the above in mind, we will create a model from basic principles to simulate Malthus’ population predictions using Microsoft Excel so that we might gain a better understanding of its implications.

**Start Simple and Increase the Complexity!**

**INITIAL HYPOTHESIS:** Population at a particular point in time is equal to the Population in the previous year plus Births.

Symbolically, we can express this in a pseudo-mathematical language that looks like:

\[
\text{Population (t)} = \text{Population (t-Δt)} + \text{Births per year} \times \Delta t
\]

INITIAL Population = 1000

Δt = Fraction of a year

Births per year = 200

So, we have a very simple model that we can compare to reality. How do they compare? Well, if we run this model using Excel, we see the linear growth of population over time. What we actually have is a model, supposedly for population growth, that matches Malthus' model for food growth.

We know that we have at least omitted a very obvious variable: Deaths! We could make our model a little more complex by adding a Death variable, represented in as a “flow” away from Population.

**IMPROVED HYPOTHESIS:** Population at a particular point in time is equal to the population in the previous year plus Births minus Deaths.

If Deaths per year is assumed to be 100, we have, symbolically:

\[
\text{Population (t)} = \text{Population (t-Δt)} + \text{Births per year} \times \Delta t - \text{Deaths per year} \times \Delta t
\]

INITIAL Population = 1000

Δt = Fraction of a year

Births per year = 200

Deaths per year = 100

However, the graph of population over time is still a linear one, just increasing at a constant absolute rate that is less than before. Can we do better?

**Add Complexity**

A key point to Malthusian growth is the idea of proportional or percentage growth, but the models thus far are assumed to have constant absolute growth rates. To introduce percentage growth, think about the
increase in population when that increase is proportional to the population level itself. In other words, the increase is dependent upon two things: The proportion and the Population.

**IMPROVED HYPOTHESIS:** Population at a particular point in time is equal to the population in the previous year plus Births minus Deaths, each of which is dependent upon the current population.

Rather than having Births determined by an “invisible hand,” we'll have it determined by a fixed proportion and by the Population itself. To do this we introduce a parameter, called Per Capita Births per year, which will be constant. Do the same sort of thing with Deaths, making it dependent upon a constant Per Capita Deaths per year and upon Population.

\[
\text{Population (t) = Population (t-Δt) + Births per year * Δt - Deaths per year * Δt}
\]

**INITIAL Population = 1000**

Δt = Fraction of a year

Births per year = Per Capita Births per year * Population

Deaths per year = Per Capita Deaths per year * Population

Now if we put in some reasonable numbers for Per Capita Births per year and Per Capita Deaths per year and run the model, what do we get?.

How does this compare to Malthus? How does it compare to our knowledge of population growth?

Well, the model pretty well captures Malthus' ideas about population growth, but it turns out not to fit well with all population data. This isn't surprising, since we haven't taken any real-world features into account, other than the propensity of populations to grow!

We'll put in one last factor to try to make the model more realistic for certain populations, leaving it to your laboratory work to take into account food, agricultural innovation, fertility rates, education, social security systems, etc. This last factor is the idea of a carrying capacity. It seems reasonable that our earth and solar system have some limit to the population it can support. If so, it has a carrying capacity, or maximum number of individuals that can survive on the planet or in a particular ecosystem.

**IMPROVED HYPOTHESIS:** Population at a particular point in time is equal to the population in the previous year plus Births minus Deaths, each of which is dependent upon the current population and a maximum population.

The previous model is extended by adding a parameter called the Carrying-Capacity, which is a (pretty big) constant. The equations for the model become:

\[
\text{Population (t) = Population (t-Δt) + Births per year * Δt - Deaths per year * Δt}
\]

**INITIAL Population = 1000**

Δt = Fraction of a year

Births per year = Per Capita Births per year * Population * (Carrying Capacity - Population)/ Carrying Capacity

Deaths per year = Per Capita Deaths per year * Population

[Note: on the term 'Per Capita Births per year' in this model that includes Carrying Capacity.]
Thus, the number of births per year is assumed in this model to be jointly proportional to the population and how close the population is to carrying capacity. When population is graphed against time it is an elongated, roughly S-Shaped curve shown below: (Population in billions, time in tens of years)

![Graph of population growth](image)

**Does the Model Explain?** Is the graph above consistent with Malthus? Other population data? What is the significance of the flat growth for the latter portions of time?

Although the graph above is not consistent with Malthus' assumption about population growth having a constant doubling time, it perhaps captures some of the spirit in the following way: There is early on an increasing rate of growth as the graphs 'curves upward.' The fact that the graph increases at smaller and smaller rates of growth as it approaches a carrying capacity is possibly what Malthus had in mind by the term 'misery.' This slow approach to a carrying capacity is perhaps the result of war, pestilence, and starvation as more and more people contend for the resources that are now at their upper bound.

What is clear is that even if the graph were a good depiction of actual world population growth, it doesn't explain much. The dynamics of population growth remain a mystery. None of the dynamic interaction of the factors related to population growth are either assumed in or deduced from the present model.

**A Note on the Idea of a Parameter.** The notion of parameter is inherent to mathematical modeling. Roughly speaking, the parameters of a model are the constants involved in the model. For example, we initially set Births per year equal to a constant, and we could have set that constant equal to anything we wanted. For that reason, Births per year was a parameter in the initial model. Once we changed the model by adding Per Capita Births per year and set Births per year equal to the product of Per Capita Births per year and Population, Births per year was no longer a parameter, but Per Capita Births per year became a new parameter, which we could set equal to something like 0.03. Introducing Carrying Capacity into the model introduced yet another parameter into the model.

Values for the parameters of a model are usually decided upon by collecting data or experimenting. However, values may be set in any way the modeler wants and the resulting model 'run' to see what the consequences are. The ability to experiment in this way is a very useful property of a mathematical model.

**The Values of Mathematical Modeling.**

1. One is forced to choose what to focus on. You must prioritize factors.
2. The modeling process helps make thoughts more precise.
3. A model helps one go beyond the surface of a phenomenon to an understanding of mechanisms and relationships.
   One can play out different scenarios, modifying assumptions, initial values, and values of parameters, to see the resulting effects.

**Problems Associated with Mathematical Modeling.**
1. The model doesn't address what you want to accomplish.
2. The model is very sensitive to initial conditions or to the values of parameters.
3. The model creates a mathematical solution to a problem that doesn't lend itself to a mathematical solution.
4. The model is too simple to mirror adequately.
5. The model is too complex to aid understanding.
6. The results are too technical to communicate.
7. The results aren't in a form that can be implemented.
8. Resources aren't adequate to implement a suggested solution.