Extended Essay

Research Question:

How does soil pH distribution affect the optimum pH of plants in an ecosystem simulation?

Subject: Biology

Word Count: 3589

Table of Contents

Introduction:

Research Question: How does soil pH distribution affect the optimum pH of plants in an ecosystem simulation?

Soil pH, the acidity of soil, can affect various processes in plants. Soil pH outside of plant's optimum range can cause denaturation of plant enzymes, disrupting necessary molecular processes[1]. Soil pH also affects the solubility of necessary minerals and nutrients in the soil. An organism's zone of tolerance is the range of values of a biotic or abiotic factor within which it can survive and reproduce[1]. In this fashion, different species of plants show varying ranges of habitable soil pH values as well as optimum soil pH values. For example, some species, such as conifer trees, demonstrate increased growth in acidic soils, while others do well in slightly alkaline soils[2].

Evolution is a change in the heritable characteristics of a population over time. The mechanism of evolution in life is natural selection[1]. Briefly, variation within a population leads to increased reproductive success in some individuals due to the possession of beneficial traits. Individuals with these characteristics generate greater number of offspring causing a higher proportion of the population with possession of these characteristics in the next generation. Overtime, this leads to a significant change in the characteristics of the population. Two types of natural selection were focused on in this report. The first, directional selection, is natural selection in which a continuously varying characteristics of a population shifts towards an extreme due to that extreme being chosen for[3]. For example, if a population of finches were introduced to a region where only hard seeds were available for consumption, the beaks of the finches may shift to become shorter and harder as they

---

[1] Andrew Allott and David Mindorff, *Biology: Course Companion* (Oxford, UK: Oxford University Press, 2014).

[2] A. Lynn Cochran, "Do Evergreen Plants Like an Alkaline or Acidic Soil?," Home Guides | SF Gate, October 7, 2016, https://homeguides.sfgate.com/evergreen-plants-like-alkaline-acidic-soil-54675.html.

[3] Paula M Mikkelsen, "Types of Natural Selection," TFG for Evolution Using Bivalves (Paleontological Research Institute, 2009), http://bivalves.teacherfriendlyguide.org/index.php?option=com_content.

enable the finches to break open more seeds. The second type of natural selection focused on is disruptive selection. Disruptive selection is natural selection in which a continuously varying characteristic of a population shifts away from a midpoint, but instead towards two extremes. For example, if the finches were introduced on an island with both hard nuts and very soft berries, finches may be selected to have either stout beaks or long delicate beaks but not medium sized beaks. Knowledge of evolution and natural selection is important because it offers a scientific explanation for how complex life may have developed. In addition, a knowledge of how populations change allows biologists to predict how current or future populations may change. For example, ecologist's might find it useful to know at what rate, a population of plants could adapt to soil pH changes caused by acid rain.

Genetic algorithms were first introduced by John Holland and are programs based off of Charles Darwin's theory of evolution by natural selection[4]. A genetic algorithm is a program in which a population of solutions to a problem is generated. Solutions which are more successful are bred and hereditary information which codes for their solutions are passed on in higher frequency to a next generation[5]. Over time, a random initial population can develop into highly optimized solutions to a problem. The problem of life is reproduction and the many solutions to it are the different forms of life. Genetic algorithms are used for a variety of purposes in mathematics, science and engineering. For example, genetic algorithms have been used by economists to model rational agents and by engineers in computer automated design and optimization.

This exploration uses a genetic algorithm to answer the research question, **"How does soil pH distribution affect the optimum pH of plants in an ecosystem simulation?"** In short, the program works by generating a population of plants with diverse optimum pH values within a habitat. Plants in habitats closer to their optimum pH are more efficient in their use of energy. Energy is used to sustain

---

[4] Vijini Mallawaarachchi, "Introduction to Genetic Algorithms - Including Example Code," March 1, 2020, https://towardsdatascience.com/introduction-to-genetic-algorithms-including-example-code-e396e98d8bf3.

[5] Atul Kumar, "Genetic Algorithms," GeeksforGeeks, August 23, 2018, https://www.geeksforgeeks.org/genetic-algorithms/.

life and reproduction. Heritable characteristics like optimum pH are passed on to progeny. The optimum pH values can then be graphed and analyzed. This simulation differs from previous simulations in several ways. First, the introduction of a limit beyond which organisms are reproductively isolated allows for speciation. This feature may be beneficial to other genetic algorithms because it has the ability to generate more than one solution to a problem in a single trial. This program also has a unique focus on the analysis of the effect of pH distributions on the first trophic level. It is hypothesized that a constant pH field distribution will lead to a unimodal normal optimum-pH distribution in the population while a habitat with split pH will lead to a bimodal optimum-pH distribution. The mechanics of this simulation are explained and justified further in the **Program Description** section.

Program Description:

The purpose of the program is to simulate the relationship between a population of phototrophic organisms and the pH of the soil in their habitat. My intent in this section is to explain and justify the mechanisms of the simulation in relation to real ecosystems.

Habitat

Excluding pH, all variables in the habitat were kept constant. For programming purposes in this simulation, the typical 0-14 pH acidity scale was scaled to a 0-100 acidity scale. This change is aesthetic and had no effect on the results. The ecosystem is simulated in a bounded rectangular habitat. An initial pH distribution could be specified by the experimenter before running the simulation. Two pH distributions were used in this experiment. First, a constant soil pH distribution was used. At any location within the habitat, the soil pH value was constant. Most real ecosystems have generally constant soil pH distributions similar to this. Second, a habitat sectioned into equally large slightly acidic and slightly basic regions was used. A divergence in the soil pH of a habitat can be found in ecosystems which have recently undergone changes due to outside factors. For example, nearby human construction may cause a change in soil pH in a given area due to the pollution of chemicals into nearby soil. Local differences in soil pH may be caused naturally through differences in the soils' clay and organic matter contents[6]. Control of the distribution of soil pH allows the experimenter to gather data on plant populations' responses to any chosen soil pH distribution.

Individuals

This simulation is an individual-based model, meaning it simulates the interactions of each organism with other organisms and its environment[7]. To do this, each organism stores data on its location, size, optimum pH etc. In real life, the genes present in plants' DNA determine their traits through the activation of corresponding metabolic pathways. As stated in the introduction, real species of plants exhibit optimum pH values at which they are more successful. Plants which are situated in areas of non-optimum soil pH tend to be less reproductively successful. Non-optimum pH can cause the denaturation of enzymes and decreased nutrient solubility[8]. These negative effects lead to decreased efficiency and less energy to be used for growth and production of offspring.

This relationship between a plant's optimum pH and their local soil pH was implemented in the simulation. Of the total possible energy produced by photosynthesis by an individual plant, it is assumed that only a percentage of this energy will actually be used. Plants which are further from their optimum pH range will waste a greater portion of their potential energy than better situated plants. To implement this, the following function was used.

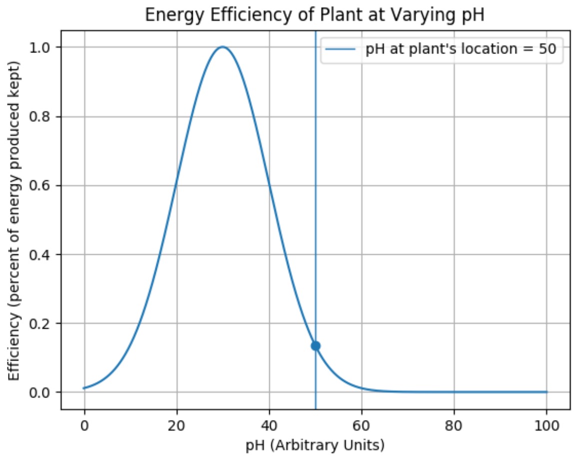
---

[6] USDA NRCS, "Soil Health, Guide for Educators," Soil Health, Guide for educators § (2014), https://www.nrcs.usda.gov/Internet/FSE_DOCUMENTS/nrcs142p2_051574.pdf.

[7] M. Khater, "EcoSim: An Ecosystem Simulation - Ecosim," Google Sites, 2012, https://sites.google.com/site/ecosimgroup/research/ecosystem-simulation.

[8] Bickelhaupt, "Soil pH: and what it means"

Fig. 1



The graph above demonstrates that a plant with optimum pH 30 units which is situated at a location of soil pH 50 units will produce about 15 % of the total energy they could possibly produce. Any plant will have a function, like this, whose maximum is centered at the plant's optimum pH. This means, that a plant will produce 100% of their total possible energy at their optimum soil pH. This function was used because it demonstrates all of the features of the relationship between optimum soil pH and local soil pH. First, soil pH will have no negative affect when close to a plants optimum pH. Second, at a specific distance, a plant will have 0 % efficiency and not be viable. In the graph, the plant's efficiency decreases to about 0% in habitats with soil pH of about 60 units or greater.

It should be noted, that the horizontal stretching of this function could determine how much pressure is put on a plant outside of their optimum pH. If the function is wide, plants will have a larger fundamental niche at which they are able to sustain life. The opposite being true if the function were horizontally shrunk. A value for this stretching was determined through preliminary testing so that it reached a point at which the initial population of plants could sustain life. The basis for variables such as this is discussed in the evaluation section of this paper.

Energy, Reproduction and Death

The simulation is run in turns. Rather than time being continuous, time is split into years. Each year, plants may produce energy, grow, reproduce, and/or die. This is similar to the cycle of seasons in temperate zones on earth, where plants have periods of growth, energy production and times for seed production.

Energy is produced based on the size of a plant. Plants produce useful chemical energy using radiant energy in photosynthesis. Photosynthetic structures like leaves are grown thin because increasing surface area allows for a larger region where a reaction can take place. In the simulation, plants are visualized as circles, through growth, plants can increase the surface area of their circle. The area of a plant's circle determines the total possible amount of energy a plant can produce. A plant receives a percentage of this energy determined by the soil pH function.

After all plants' energies have been calculated, death, growth and seed production are calculated. All organisms use energy to sustain life. For the plants, it is assumed that a larger amount of energy will be required to sustain the life of larger plants. In the simulation, the value calculated for the required energy to sustain life is proportionate to the surface area of a plant. Plants who cannot produce enough energy to reach this value, die. In addition, plants which have reached an age limit (10 years in the simulation) die of old age. Plants with surplus energy can use this energy for growth or seed production. Plants in the simulation have a growth limit gene. This gene determines the limit to which the radius of a plant will grow.

After all plants which can grow have grown, plants with additional energy can use this energy for sexual reproduction. The plants simulated are dioecious, producing flowers of only one sex[9]. For each pollen produced by a male, a list of viable females is produced. A random female is chosen from this list to produce a seed. Each gene of the gamete is selected from a normal probability density

---

[9] John Chau, "Plant Morphology: Vegetative and Reproductive," University of Washington, 2016, http://courses.washington.edu/bot113/summer/LectNotes/2011/Lecture1_2.pdf.

function centered at the mean of the parents' corresponding genes. This simulates the randomness produced through crossing over and mutation. The location of seeds is generated randomly.

At a specific genetic difference, plants are reproductively isolated (Discussed further in Evaluation). In real life, reproductive isolation of two sexually reproducing organisms can be due to physical separation, behavioral separation, an inability to form a gamete or the production of non-viable or sterile offspring[10]. Reproductive isolation is a necessary ingredient for speciation. The optimum pH of a plant is determined by many genes. In this simulation, it is assumed that these genes can be modeled using a single value for optimum pH. From this, it is reasoned that simulated plants which have distant optimum soil pH values will have significant differences in their genetic sequences. This difference will cause their offspring to be non-viable. A value for a distance of reproductive isolation was determined through preliminary testing.

Randomness and Limiting Factors

Randomness and limiting factors are essential components of evolution. Without variation or mutation, a population has no means for change. Randomness is also important in testing for unexpected weaknesses in a model. The genes and locations of the initial population of plants are produced randomly. Randomness also takes a role in how genes are passed on, which mates are chosen and the location of new seeds. Limiting factors are constraints on populations which limit growth. They are important in evolution because they act as incentives for improvement. In this simulation, the size of the habitat, the age limit and the energy required to sustain life are limiting factors.

---

[10] Roy Caldwell, "Reproductive Isolation," Understanding Evolution, 2009, https://evolution.berkeley.edu/evolibrary/article/evo_44.

Evaluation:

Justification:

The main intent of the program is to answer the research question, “How does soil pH distribution affect the optimum pH of plants in an ecosystem simulation?” That the program runs consistently supports its validity. In addition, the program is built off of sound principles and recorded phenomena that underly how ecosystems normally run.

In addition to answering the research question, this program may also be useful for understanding or making predictions concerning changes in the soil pH of a habitat. It allows for future predictions on the effect of pH shifts which it wasn’t designed to measure. For example, habitats with soil a range in pH, constantly shifting pH, random pH distributions, etc. could all be simulated. The ability to make accurate future predictions defines a useful scientific model. This feature of the program could be helpful to conservationists hoping to predict the effect of a change of pH on species distribution or the abundance or diversity of an area. Another benefit of the use of a simulation in comparison with a field study is that the only limiting factors in sample size with a program like this are time and computational power.

Although not its intended use, the program’s simplicity allows for its use in educational purposes. Students of biology may benefit from an interactable and visual depiction of natural selection and speciation.

Potential Improvements:

One of the areas in which this study could be improved is in the quantification of uncertainty. Schuwirth et al. give the quantification of uncertainty as one of six requirements for useful ecological models[11]. One of the difficulties in designing a simulation is in evaluating its accuracy. While the quality of the simulation has been argued through its relation, the author has found no means with which to quantify the uncertainty in the results. This is one weakness of simulating an ecosystem rather than collecting data to form a conclusion.

This simulation focused on the relation between plants and soil pH. To do this some complexity in real ecosystems has been simplified. This is justified, as all models require the simplification of the phenomena they model through keeping all other variables constant. The relative simplicity of the interactions of organisms in this simulation allows population changes in optimum soil pH to be unhindered and clearly demonstrates any relationship found to be causal. However, there may have been important factors which are present in the real complex ecosystems which have gone missing in this process of simplification. For example, factors such as nutrient availability, weather and the seasons may affect interactions in real ecosystems. Although, randomness was incorporated into the simulation during the mixing of genetic information, the placement of seeds and the creation of the initial population, to improve the simulation, random events such as weather events or predator interactions could be included as well.

To model a specific real-life population, the simulation's variables must be changed. Variables such as initial population size and the energy cost to produce seeds were set after preliminary testing revealed life sustaining ranges for these variables. Any application of the simulation to a real ecosystem would have to find exact corresponding variable which match the population and ecosystem. This paper does not pose a method of calculating these variables from real life plant populations, but this could be tackled in future research.

One variable in particular, the speciation limit, creates several weaknesses. In real life, reproductive isolation of two sexually reproducing organisms can be due to physical separation, an inability to form a gamete or the production of non-viable or sterile offspring. What determines

---

[11] Nele Schuwirth et al., "How to Make Ecological Models Useful for Environmental Management," *Ecological Modelling* 411 (2019): p. 108784, https://doi.org/10.1016/j.ecolmodel.2019.108784.

whether an offspring is non-viable or sterile may vary greatly and depend on other variables such as plant size or fertilization mechanisms. Some of this complexity may be missing in the simulation.

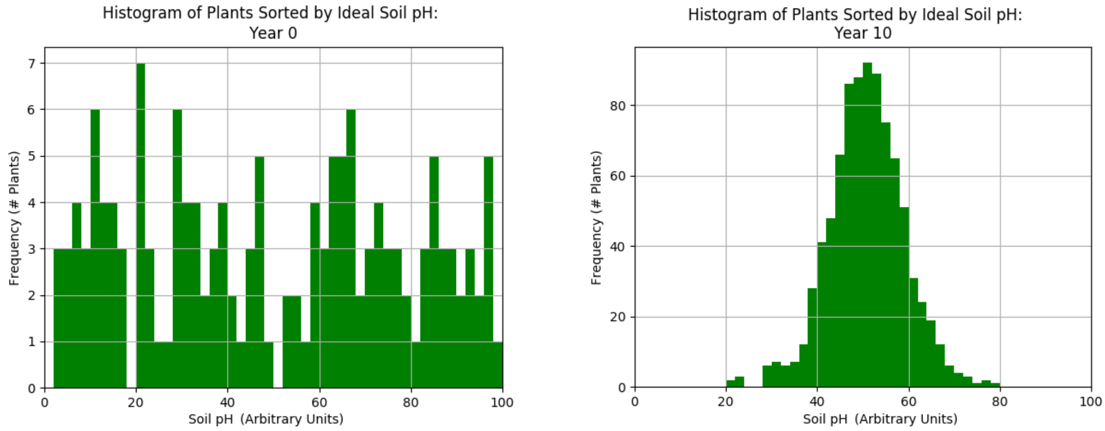
Lastly, plants' relationships with soil pH may be more complex than demonstrated in the model. Some plants may cause small shifts in soil pH. For example, the needles of conifer trees are slightly acidic. A large quantity of these needles could shift the soil pH to become more acidic[12]. Although not all species have effects like this on soil pH, some augmentations could be made to the program if one wished to simulate these plants.

[12] Emma Erler, "Do Pine Trees and Pine Needles Make Soil More Acidic?," UNH Extension, October 16, 2019, https://extension.unh.edu/blog/do-pine-trees-and-pine-needles-make-soil-more-acidic.

Conclusion:

Two soil pH distributions were inputted into the simulation, a constant distribution and a split distribution (see Fig. 1). Example results of the constant distribution are shown below

Fig. 2



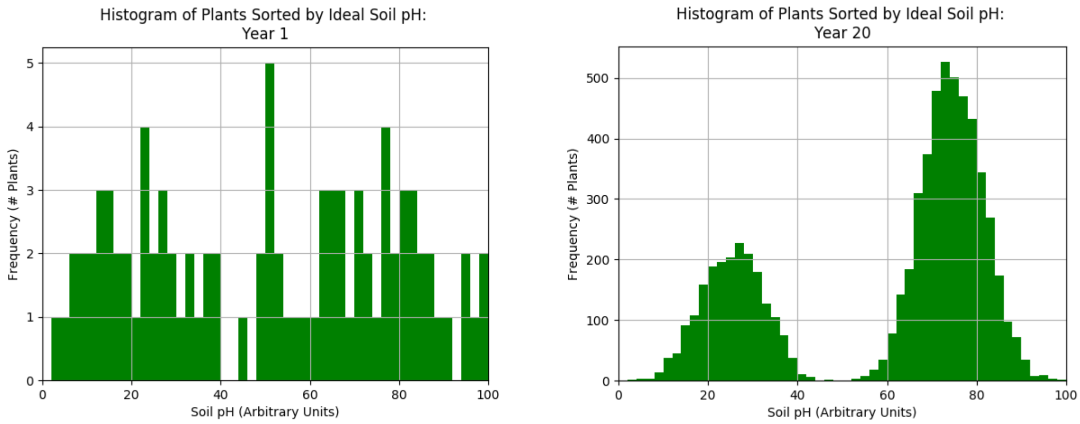
The results support the hypothesis that a constant soil pH distribution would develop a normally distributed population with regards to the optimum pH trait. The first figure shows an initial population with random optimum soil pH distribution. For most trials, by year 2, organisms with extreme optimum soil moisture values died out. The most successful plants continued to reproduce. Due to death by old age, after 10 years, none of the initial population remained. At this time, the optimum soil pH distribution of the population resembled a normal distribution centered around a pH of 50 units, the same pH as the habitat. The simulation was run multiple times. Each trial, a similar normal distribution was reached. This matches normal distributions found in other studies of natural selection such as Abrahamson et al's study of selection pressures on gall flies[13]. This link to the

---

[13] Warren G. Abrahamson et al., "Variation in Selection Pressures on the Goldenrod Gall Fly and the Competitive Interactions of Its Natural Enemies," *Oecologia* 79, no. 1 (1989): pp. 15-22, https://doi.org/10.1007/bf00378234.

literature supports the validity of the simulation and the above relationship while reinforcing the validity of the theory of natural selection.

Fig. 3



The results support the hypothesis that a split soil pH distribution would develop a bimodal distribution of optimum soil pH values. As before, this simulation started with a random distribution. However, a difference is that the bimodal distribution required a longer amount of time to reach near constancy than a unimodal distribution. For most trials, the population began to stabilize 15-20 years into the simulation. In comparison, the unimodal distribution most often stabilized in around 10 years. This may have been due to reproduction between organisms closer to the center on either side of the distribution. This would have produced new plants which died shortly after being created but disrupted the distribution. The inequality in the magnitude of the modes in the bimodal distribution was random and only seen in the trial shown. The distribution in Figure 3 also matches graphs of disruptive selection found in the literature. This further supports the validity of the simulation in predicting the effect of soil pH distributions on a population.

As stated above, in response to the research question, “How does soil pH distribution affect the optimum pH of plants in an ecosystem simulation?”, these results support the hypotheses that a constant soil pH distribution would develop a normally distributed population with regards to the optimum pH trait and that a split soil pH distribution would develop a bimodal distribution of optimum soil pH values.

Further Research

There are several ways this research could be expanded upon. Future research should address the improvements mentioned in the "Evaluation" section of this report. To start, researchers could customize variables to match specific plants or regions. As mentioned in the "Justification" portion of this report, the simulation allows for any soil pH distribution as input. Future research may use this feature to investigate the effect of real habitat changes on populations. In addition, the variables of the initial population of plants such as optimum soil pH could be generated to fit known variables of a species of plant to predict its reaction to a change in habitat.

Although unnecessary for this simulation, future research could incorporate other trophic levels in the program. For example, simulating herbivores, carnivores and/or decomposers would add a new level of complexity to this ecosystem which could result in emergent properties such as special behaviors or competition. However, modeling a full ecosystem on the same scale as this program would require increased processing capabilities.

Bibliography

Abrahamson, Warren G., Joan F. Sattler, Kenneth D. Mccrea, and Arthur E. Weis. "Variation in Selection Pressures on the Goldenrod Gall Fly and the Competitive Interactions of Its Natural Enemies." *Oecologia* 79, no. 1 (1989): 15–22. https://doi.org/10.1007/bf00378234.

Allott, Andrew, and David Mindorff. *Biology: Course Companion*. Oxford, UK: Oxford University Press, 2014.

Bickelhaupt, Donald. "Soil PH: What It Means." SUNY college: ESF, 2020. https://www.esf.edu/pubprog/brochure/soilph/soilph.htm.

Caldwell, Roy. "Reproductive Isolation." Understanding Evolution, 2009. https://evolution.berkeley.edu/evolibrary/article/evo_44.

Chau, John. "Plant Morphology: Vegetative and Reproductive." University of Washington, 2016. http://courses.washington.edu/bot113/summer/LectNotes/2011/Lecture1_2.pdf.

Cochran, A. Lynn. "Do Evergreen Plants Like an Alkaline or Acidic Soil?" Home Guides | SF Gate, October 7, 2016. https://homeguides.sfgate.com/evergreen-plants-like-alkaline-acidic-soil-54675.html.

Erler, Emma. "Do Pine Trees and Pine Needles Make Soil More Acidic?" UNH Extension, October 16, 2019. https://extension.unh.edu/blog/do-pine-trees-and-pine-needles-make-soil-more-acidic.

Khater, M. "EcoSim: An Ecosystem Simulation - Ecosim." Google Sites, 2012. https://sites.google.com/site/ecosimgroup/research/ecosystem-simulation.

Kumar, Atul. "Genetic Algorithms." GeeksforGeeks, August 23, 2018. https://www.geeksforgeeks.org/genetic-algorithms/.

Mallawaarachchi, Vijini. "Introduction to Genetic Algorithms - Including Example Code," March 1, 2020. https://towardsdatascience.com/introduction-to-genetic-algorithms-including-example-code-e396e98d8bf3.

Mikkelsen, Paula M. "Types of Natural Selection." TFG for Evolution Using Bivalves. Paleontological Research Institute, 2009. http://bivalves.teacherfriendlyguide.org/index.php?option=com_content.

NRCS, USDA, Soil Health, Guide for educators § (2014). https://www.nrcs.usda.gov/Internet/FSE_DOCUMENTS/nrcs142p2_051574.pdf.

Schuwirth, Nele, Florian Borgwardt, Sami Domisch, Martin Friedrichs, Mira Kattwinkel, David Kneis, Mathias Kuemmerlen, Simone D. Langhans, Javier Martínez-López, and Peter Vermeiren. "How to Make Ecological Models Useful for Environmental Management." *Ecological Modelling* 411 (2019): 108784. https://doi.org/10.1016/j.ecolmodel.2019.108784.