JOURNAL OF
EMERGING INVESTIGATORS

VOLUME 3, ISSUE 11 | NOVEMBER 2020
emerginginvestigators.org

Reaching new heights
Investigating the thermodynamics of hypersonic speeds in the stratosphere

Diagnosing Parkinson's disease
A new method using voice data from patients

Diophantine density
Patterns abound in Diophantine equation solutions

Fighting climate change with algae
Optimizing the biofixation process to reduce atmospheric CO_2 levels

Clean rice for all
An inexpensive technique for removing arsenic from rice

JOURNAL OF EMERGING INVESTIGATORS

The Journal of Emerging Investigators is an open-access journal that publishes original research in the biological and physical sciences that is written by middle and high school students. JEI provides students, under the guidance of a teacher or advisor, the opportunity to submit and gain feedback on original research and to publish their findings in a peer-reviewed scientific journal. Because grade-school students often lack access to formal research institutions, we expect that the work submitted by students may come from classroom-based projects, science fair projects, or other forms of mentor-supervised research.

JEI is a non-profit group run and operated by graduate students, postdoctoral fellows, and professors across the United States.

Contents

VOLUME 3, ISSUE 11 | NOVEMBER 2020

JOURNAL OF
EMERGING INVESTIGATORS

Article

Analysis of the effects of positive ions and boundary layer temperature at various hypersonic speeds on boundary layer density

Akash Selvakumar[1], Sridhar Palanisamy[2]
[1]North Creek High School, Bothell, Washington, USA
[2]Karpagam University, Coimbatore, India

SUMMARY

Vehicles traveling at hypersonic speeds ionize the air around them. This ionization leads to the possibility of utilizing an induced positive surface charge to lower the density of the boundary layer flow electrostatically. Reducing the boundary layer's density would have several positive effects, such as lowering the drag and heat transfer. The study's goal was to identify the Mach numbers for which the electrostatic drag and heat transfer manipulation would be most applicable inside the stratosphere. We hypothesized that the potential to use induced positive surface charge repulsion to lower boundary layer density would increase with speed, and there will be greater (>10%) potential to decrease boundary layer density at a lower bound estimate of Mach 18-19. We also explored the extent to which extrapolation through a quadratic model based on Dr. Jesse R. Maxwell's data in *Morphing Waveriders for Atmospheric Entry* could be used to approximate maximum boundary layer temperature values. The experiments were conducted using computational fluid dynamics software. The study demonstrated that, on average, higher Mach speeds resulted in a considerably higher potential decrease in density. Also, as predicted by the hypothesis, a significant potential reduction in density occurred near Mach 19. The study also supported using the quadratic model based on Maxwell's data for approximating the maximum boundary layer temperature up to at least Mach 26. This study highlights that further research on the surface charge method is warranted as we seek to explore higher hypersonic speeds within the stratosphere.

INTRODUCTION

The desire to achieve higher speeds is a constant fascination with humans. As we delve into the hypersonic spectrum, there is a constant search for ways to further increase the speed of our atmospheric vehicles. In order to do so, it is essential we identify methods to reduce drag and lower the convective heat transfer(heat transfer from the surrounding hot air to the vehicle), which can be fatal to vehicle bodies located inside the extreme temperatures present in hypersonic airflows. One way to achieve a reduction in skin drag and convective heat transfer is through decreasing the density of the boundary layer. Decreasing the density of the boundary layer leads to a reduction in skin friction and thus a reduction in skin drag (1). Convective heat transfer can be considered as the combined effects of conduction and fluid motion (advective) heat transfer (2). The heat flux φ_q for advective heat transfer is modelled by the equation:

$$\varphi_q = v\rho c_p \Delta T$$

where ρ is density, c_p is heat capacity at constant pressure, ΔT is the difference in temperature, and v is velocity. Thus, changing the density of flow inside the boundary layer (near the surface of the vehicle) the advective and thus convective heat transfer is decreased.

Our study considered the induced surface charge repulsion method for lowering boundary layer density (3). Inducing positive surface charge on the vehicle will electrostatically repel positive ions and thus lower boundary layer density (3). The maximum potential reduction in boundary layer density through induced surface charge, which will be calculated for specified trails in the experiment, will thus be equivalent to the percentage of the density that can be attributed to positive ions.

We performed a study to analyze the maximum potential reduction in the boundary layer possible by positive surface charge repulsion at various speeds. The study utilized the Ansys Fluent Computational Fluid Dynamic Software to perform several trials at different Mach numbers (a dimensionless measurement of speed with respect to the local speed of sound) and gather average boundary layer static temperature data. We repeated each trial twice and averaged the results to produce reliable data. This data was then analyzed to answer questions regarding the relationship between Mach number and maximum potential density decrease by induced positive surface charge repulsion. Furthermore, we analyzed the data to validate a proposed model and understand the relationship between boundary layer temperature and Mach number. The method of induced

charge has a primary application in the high hypersonic spectrum because the air is not at high enough temperatures to be ionized at subsonic, supersonic, and low hypersonic speed ranges in regular air conditions, rendering the effect of electrostatically repulsion negligible (4-5).

Temperature is a measurement of the average kinetic energy(6). When a vehicle moves through air it transfers some of its kinetic energy to the particles around it via collisions (7). The amount of kinetic energy gained by the particles is proportional to the kinetic energy of the vehicle. Therefore, air particles around vehicles traveling at higher velocities gain more kinetic energy and temperature (7). Since increasing temperature of the air results in increased ionization of the air particles (4), we hypothesized that the potential for the induced positive surface charge method to lower boundary layer density will increase with speed. Upon performing our CFD study, we demonstrated that, on average, higher Mach speeds resulted in a considerably higher potential decrease in density supporting this hypothesis.

RESULTS

Preliminary Estimate

A study done by Maxwell provided the maximum temperature values inside the boundary layer for vehicles traveling inside the stratosphere at speeds up to Mach 10 (5). To estimate maximum boundary layer temperature values at Mach numbers greater than 10, we performed quadratic regression. From the quadratic regression, we determined the quadratic best fit function for Maxwell's data to be:

$$T(s) = 196 + 43.9s + 23.8s^2$$

where ***T*** is the maximum boundary layer temperature and ***s*** is the Mach number (**Figure 1**).

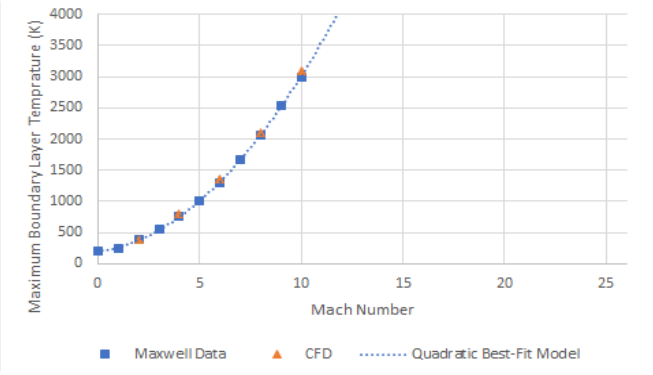

Figure 1. Comparison of published experimental and CFD maximum boundary temperature values. The Quadratic Best fit model (dotted blue line) based on Maxwell's published maximum boundary layer data (5) is included to show extrapolation to higher values. The Graph illustrates the closeness between the Quadratic model estimates, Maxwell's maximum boundary layer temperature values (blue squares), and CFD maximum boundary temperature values (orange triangles).

Table 1. Comparison of CFD estimates, published experimental data, and quadratic model predictions of maximum boundary layer temperature at lower Mach numbers.

Mach Number	CFD Maximum Boundary Layer Temperature Values (K)	Maxwell Maximum Boundary Layer Temperature Values (K)	Quadratic Best Fit Model Maximum Boundary Layer Temperature Estimates (K)	Percent Deviation of CFD Maximum Boundary Layer Temperature Values from Maxwell Values (%)	Percent Deviation of Quadratic Model Maximum Boundary Layer Temperature Estimates from Maxwell Values (%)
2	390.16	380	379.0	2.67	0.26
4	794.34	770	752.4	3.16	2.29
6	1359.79	1300	1316.2	4.60	1.25
8	2096.45	2070	2070.4	1.28	0.02
10	3097.96	3000	3015.0	3.27	0.50

This quadratic model closely adheres to Maxwell's data (**Table 1**). From previous Computational Fluid Dynamics (CFD) static temperature simulation results, we observed that a significant portion of the boundary layer is near the maximum boundary layer temperature (5). By using this quadratic model for the maximum temperature inside the boundary layer (**Figure 1**), we can also hypothesize that a lower-bound approximate value of Mach 18-19 from which we will see a significant (For our study we are defining 10% and above as significant since our educated assumption is that this level of reduction will produce measurable effects in drag and convective heat transfer reduction) decrease in drag and convective heat transfer.

As predicted by this hypothesis, we ascertained that a significant (10% or greater) potential reduction in density occurred near Mach 19. Furthermore, the study provided reasonable support for using the above quadratic best fit function for estimating maximum boundary layer temperature.

CFD Study

The potential reduction in boundary layer density through positive surface charge and its variance in regard to vehicle speed was determined by first identifying the average temperature of the boundary layer at various Mach speeds. The average boundary layer temperature was identified through utilizing the Ansys Fluent CFD software to simulate the vehicle moving at various speeds at an altitude of 25 km above sea level and for each Mach number simulation we selected 20 points for temperature reading evenly spaced throughout the boundary layer to get a good measurement for the average temperature within the boundary layer (**Figure 2**). The average temperature data was further used to determine the percent decrease in density potentially achievable by induced surface charge repulsion for each trial.

Furthermore, the quadratic model for maximum boundary layer temperature was validated by using the CFD simulation to identify the point of maximum boundary layer temperature. This point was identified by taking the temperature at the surface of the vehicle towards the tip of the cone. Preliminary CFD trials were performed at Mach 2, 4, 6, 8, and 10. The CFD maximum boundary layer temperature data showed strong adherence with previous experimentation published in Maxwell's study (**Table 1**), which strengthened the validity of

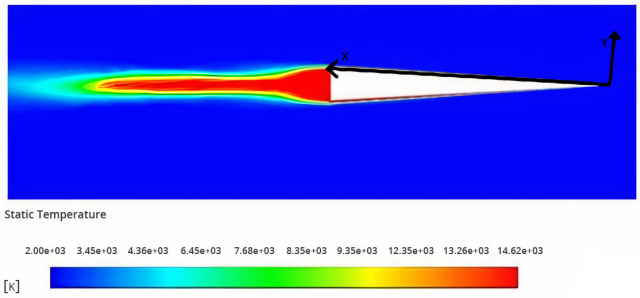

Figure 2. CFD Static Temperature Data for Mach 24. The x and y axis on the CFD temperature diagram show the position axis used to determine points for data collection. The diagram also shows the static temperature values (K) for the air around the vehicle.

this study. Afterwards, trials were performed for Mach 12, 14, 16, 18, 20, 22, 24, and 26. Additional trials were conducted at Mach 19, 21, 23 to understand the data better and identify a more exact Mach number for significant air drag reduction. The potential boundary layer density decrease was not calculated for Mach 14 and lower because the ionization of air is negligible for average temperature values lower than 6000 K (7).

On average, increasing Mach number resulted in a significant increase in average boundary layer temperature and a considerable potential decrease in boundary layer density (**Table 2**). The data also indicates that the lowest Mach number with a significant (10% or greater) potential reduction in boundary layer density occurs at slightly less than Mach 19. Thus, based on the CFD data, the lowest whole Mach number with a significant potential decrease in density will be Mach 19.

To validate the quadratic model, a mean percent error (MPE) of 3.209% was seen for the extrapolation model values compared to the actual data for maximum boundary layer temperature gathered by the CFD (**Table 3**). Furthermore, the average percentage deviation of the Maximum Boundary layer Temperature from Average Boundary Layer Temperature (**Table 3)** was 7.363%.

DISCUSSION

The study was performed to understand the relationship between Mach number and boundary layer temperature and ion density. This was used to determine the potential applicability of induced positive surface charge repulsion to lower boundary layer density. The results of this study's experimentation were supported by the strong adherence of the study's CFD determined maximum boundary layer temperature values to previously determined maximum boundary layer temperature values (5) (**Table 1**). The trials conducted supported the hypothesis that the potential to use induced positive surface charge repulsion to lower boundary layer density would increase with speed, and there will be significant (>10% density decrease) potential to decrease boundary layer density at a lower bound estimate of Mach 18-19. From the results, we can observe that increasing the Mach number results in a greater potential to decrease boundary layer density using induced surface charge. The data also showed that a greater than 10 percent potential decrease in density occurring at Mach 19. Thus, we can ascertain that at speeds greater than or equal to Mach 19, inside the stratosphere, there will be significant potential to decrease boundary layer density through repelling the positive ions in the boundary layer. Speeds of upwards Mach 9 have already been reached inside the stratosphere (8), while speeds of up to Mach 17 - 22 have been reached by aircraft at higher altitudes (9). Predictions for scramjet-powered aircraft, which operate inside the stratosphere, indicate that speeds of Mach 20 and above are achievable (10). Therefore, while this positive surface charge method is not currently applicable for stratosphere flight, this may change in the future. For instance, the induced surface charge method may have application in future military hypersonic vehicles and weapon systems traveling at very high hypersonic speeds that seek to operate at shallower trajectories inside the lower atmosphere (stratosphere) to avoid interception.

The quadratic model's predictions for maximum boundary layer temperature values at higher Mach numbers seem to be closely correlated with CFD reading. This indicates

Table 2. Data and results for determination of maximum potential percentage reduction in boundary layer density by positive surface charge at various Mach numbers.

Mach Number	Average Boundary Layer Temperature (K)	Percent Potential Decrease in Density (%)
16	6501.89	0.19
18	8469.74	4.22
19	9348.07	12.05
20	10347.45	33.03
21	11257.44	56.76
22	11715.59	64.15
23	13259.22	91.83
24	13818.99	95.71
26	17958.88	99.85

Table 3. Data for validation of quadratic model for maximum boundary layer temperature and analysis of Boundary Layer temperature

Mach Number	CFD Average Boundary Layer Temperature (K)	CFD Maximum Boundary Layer Temperature (K)	Quadratic Model Values for Maximum Boundary Layer Temperature (K)	Percent Error for Quadratic Model Results and CFD Maximum Boundary layer Temperature (%)	Percent Deviation of Maximum Boundary Layer Temperature from Average Boundary Layer Temperature (%)
2	354.16	390.16	379.00	2.86036	10.16505
4	734.36	794.34	752.40	5.27985	8.16766
6	1269.66	1359.79	1316.20	3.20564	7.09913
8	1942.86	2096.45	2070.40	1.24258	7.90549
10	2943.34	3097.96	3015.00	2.67789	5.25309
12	4118.32	4341.95	4150.00	4.42082	5.43005
14	5229.6185	5766.92	5475.40	5.05504	10.2742
16	6501.8925	7066.22	6991.20	1.06167	8.67943
18	8469.74	8998.36	8697.40	3.34461	6.24123
19	9348.07	9898.16	9621.90	2.79102	5.88451
20	10347.45	10921.80	10594.00	3.00134	5.55068
21	11257.44	11956.32	11613.70	2.86560	6.20813
22	11715.59	13180.35	12681.00	3.78859	12.50266
23	13259.22	13956.33	13795.90	1.14951	5.25755
24	13818.99	14876.24	14958.40	0.55229	7.65070
26	17958.88	18952.57	17426.20	8.05363	5.53314

that this quadratic model may be a reasonable method for approximating the maximum boundary temperature up to at least Mach 26. Additionally, the percent deviation of the maximum boundary layer temperature from the average boundary layer temperature seems to validate the observation that a significant portion of the boundary layer is near maximum boundary layer temperature. On average, a low standard deviation of less than 10 percent was observed. This observation highlights the possibility that maximum boundary layer temperature can be used as a rough estimate for the average boundary layer temperature at high Mach numbers. However, further study and additional trials would be necessary to validate this claim thoroughly.

A primary cause for fluctuations in the data can be attributed to the number of iterations that were run for each simulation. The higher the iterations the more accurate the CFD data; though $5x10^5$ iterations were performed, there may still be some inaccuracies thus resulting in fluctuations in the data. There could also be a slight variation in the placement of the cursor when recording data points. This variation would primarily serve to affect the maximum boundary layer temperature readings from the CFD, where the cursor must be located at the tip of the cone near the surface (leading edge) (11), by giving data points lower than the actual maximum. Finally, some errors in determining the maximum potential density decrease may result from using the concentration graph. The graph provides reasonably accurate estimates based on previous studies, and the manual identification process may result in some error.

It is important to note that the study used the average boundary layer temperature to get the percent density of the mixture that can be attributed to positive ions. This method would only serve to give an approximate answer. A more accurate means of identifying the density percentage is by calculating the percent density attributed to positive ions at each point and integrating through the entirety of the boundary layer. Since this operation would be impractical with the method of density identification used for this project, it was not used. However, if coupled with another percent density identification method, such as the Saha ionization equation method, this method could determine even more accurate data for future research. Furthermore, in future studies using CFDs, more accurate density change approximation may be achieved with the Saha ionization equation. This equation becomes applicable for relatively low ionization plasmas, as is the case with the air inside the boundary layer. When solved as a system of equations alongside a simple density relationship, the Saha ionization equation could better estimate the maximum density change possible at a given temperature and boundary layer density.

Further research with wind tunnels and the physically inducing charge would substantially contribute to this experiment. Physical experiments would allow us to examine a wide variety of data and thus would help get quantitative answers to questions regarding the effect of positive surface charge on air drag and convective heat transfer. Physical experiments will also be the best approach to identify the amount of positive surface charge needed to achieve significant effects and allow for research into various charge distributions to optimize effects. For example, further physical research can be performed to optimize charge distributions to optimize the decrease in heat transfer while lowering the decrease in drag for application of induced charge during reentry of spacecraft. Additional research with CFDs could consider the potential effect of positive surface charge in reducing the air density in front of the vehicle. Using a positive surface charge to reduce the perceived freestream density of the flow would have significant and already quantified effects on drag and convective heat transfer (12). However, since the temperature in front of the vehicle is considerably lower than the temperatures inside the boundary layer, much higher speeds would be necessary to produce the same density reduction. However, this research could have applications for future extremely fast hypersonic vehicles.

METHODS

Computational Fluid Dynamic Analysis

Solidworks was used to create a mesh of the hypersonic vehicle for the simulation. The vehicle was modeled by a cone, which is an effective model for hypersonic vehicle bodies (13). The cone was created to the specified dimensions of 5 m length and 0.5 m radius. The objects properties were defined as solid. The Ansys Fluent CFD environment was used to model the temperature of the air inside the boundary layer. The Solidworks mesh of the hypersonic vehicle was imported into the CFD. We configured the CFD simulation as Hypersonic flow at an altitude (25 km above sea level) and set the specified Mach number for the trial. Each trial was set to have 5x10^5 iterations performed to get accurate results. The simulation was set to return the velocity and static temperature data. Each trial was repeated twice with collected data averaged to reduce error.

Trials were held at Mach Numbers 2, 4, 6, 8, 10, 12, 14, 16, 18, 19, 20, 21, 22, 23, 24, 26. The trials at 2, 4, 6, 8, 10 served as preliminary trials to verify the accuracy of our CFD simulations. The velocity data was used to approximate the width of the boundary layer (**Figure 3**). The boundary layer is the area in which the velocity of the flow is changed by

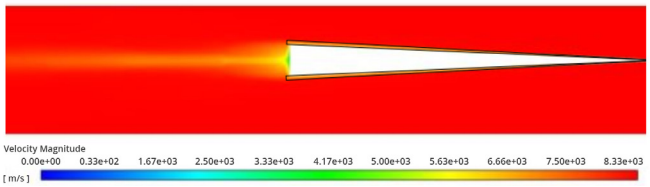

Figure 3. CFD Velocity Data for Mach 24. The boundary layer is outlined in black on a CFD velocity diagram. The diagram shows the velocity of the air (m/s) around the vehicle relative to the vehicle.

the viscous interactions (14). Thus, the areas in the data around the sides of the vehicle where the velocity changed thus indicated the boundary layer. An outline of the boundary layer was made and overlaid on the static temperature data reading indicating the whereabouts of the boundary layer for accurate data collection.

Average Boundary Temperature Identification

With the boundary layer marked, 20 points were chosen by hand evenly within the boundary. Due to symmetry, the points were taken at the upper boundary layer (**Figure 3**). We set the position axis with the x axis as distance in meters from the origin located at the tip of the vehicle and the y axis as distance in meters from the vehicle (**Figure 2**). The points were chosen at roughly 0.0, 0.5, 1.0, 1.5, 2.0, 2.5, 3.0, 3.5, 4.0, 4.5, 5.0, 5.5 as x values and y values varying evenly within the boundary layer. The number of points chosen at each value vary proportionately with the thickness of the boundary layer at each x value. Taking an average of the static temperature values at these points was used to produce a good estimate for the average temperature within the boundary layer.

Maximum Boundary Layer Temperature Identification and Calculations

The maximum boundary layer temperature was identified by taking the value at the point (0,0) utilizing the position axis defined above (**Figure 2**). This gives the value at the leading edge and tip of the cone where the temperature is the highest (11). Percent error calculation of quadratic model data for CFD maximum boundary layer temperature was calculated by the equation:

$$\%e = |T_m - T_c|/T_c \times 100\%$$

where T_m is the maximum boundary layer temperature calculated from the quadratic model and T_c is the maximum boundary layer temperature obtained from the CFD simulation. The percent deviation of maximum boundary layer temperature from average boundary layer temperature was calculated similarly using the equation:

$$\delta = |T_z - T_c|/T_a \times 100\%$$

where T_a is the average boundary layer temperature obtained from the CFD simulation and T_c is the maximum boundary layer temperature obtained from the CFD simulation. Similar calculations were performed for the percent deviation of CFD simulation maximum boundary layer temperature values and Quadratic Model maximum boundary layer temperature estimates from Maxwell's maximum boundary layer temperature values.

Maximum Potential Percent Decrease in Boundary Layer Density Calculation

The concentration of positive ions inside the boundary layer was determined based on a temperature concentration graph provided by Professor J. E. Shepherd for air, based on experimental data (**Table 2**) (4). This graph was utilized to get a rough estimate for the percentage of the density attributed to positive ionic species. The maximum potential percent decrease in density by positive surface charge repulsion is equivalent to the percentage of the density attributed to positive ions. Thus, the maximum potential percent decrease in density by positive surface charge repulsion is ($\%\Delta\rho$) was thus identified by the equation:

$$\%\Delta\rho = (\Sigma m_n C_i)/(\Sigma m_n C_n + \Sigma m_i C_i)$$

where $\Sigma m_i C_i$ is the summation across all positive ions of the molar mass of each positive ion multiplied by the number concentration of each ion, $\Sigma m_n C_n$ is the summation across all species that are not positive ions of the molar mass of the species multiplied by the number density of the species.

ACKNOWLEDGMENTS

I would like to thank my parents for supporting me through the research, paper writing, and submission process. I am very grateful for Professor J. E. Shepherd for taking the time to provide me with the necessary information to calculate ion concentration. I would also like to thank Mr. Henning for helping me through some of the complex math when I was gaining preliminary understanding of the topic. Finally, I would like to thank Ms. Haupt for teaching the data analysis and scientific paper writing skills that were extremely helpful for this research.

Received: July 11, 2020
Accepted: September 24, 2020
Published: October 2, 2020

REFERENCES

1. Suraweera, M. V., Mee, D. J., & Stalker, R. J. "Skin Friction Reduction in Hypersonic Turbulent Flow by Boundary Layer Combustion." American Institute of Aeronautics and Astronautics, 2005, pp. 9-10.
2. "Convection Heat Transfer." *McGraw-Hill Higher Education*, McGrawHill Education, www.mhhe.com/engcs/mech/cengel/notes/ConvectionHeatTransfer.html.
3. Tajmar, Martin. *Advanced Space Propulsion Systems.* Springer-Verlag, 2003.
4. Shepherd, J. E. "Re: Request for information/guidance - Sub: Saha ionization" Received by Akash Selvakumar, Jun 18, 2020. E-mail Communication.
5. Maxwell, Jesse R. *Morphing Waveriders for Atmospheric entry*. 2019. University of Maryland, PhD dissertation.

DRUM, https://drum.lib.umd.edu/handle/1903/21909.
6. "Temperature and Kinetic Energy." Chemistry 301, University of Texas, ch301.cm.utexas.edu/section2.php?target=gases%2Fkmt%2Ftemp-kinetic-energy.html.
7. Atluri, S. N., *et al*. *Computational Mechanics '95 Volume 1 and Volume 2 Theory and Applications*. Springer Berlin, 2014.
8. Martin, Guy. "Hypersonic Pioneer: The X-43A." Aircraft Information, www.aircraftinformation.info/art_x43.htm.
9. Acton, James M. "Hypersonic Boost-Glide Weapons." Routledge Taylor & Francis Group, 2015, pp. 203–204, scienceandglobalsecurity.org/archive/2015/09/hypersonic_boost-glide_weapons.html.
10. McClinton, Charles. "X-43: Scramjet Power Breaks the Hypersonic Barrier." Dryden Lecture. 44th AIAA Aerospace Sciences Meeting and Exhibit, 9 Jan. 2006, Reno, Nevada, web.archive.org/web/20110724231440/http://www.aiaa.org/Participate/Uploads/AIAA_DL_McClinton.pdf.
11. Kasen, Scott D. *Thermal Management at Hypersonic Leading Edges*. 2013. University of Virginia, PhD dissertation
12. Hanquist, Kyle M., and Iain D. Boyd. "Plasma Assisted Cooling of Hot Surfaces on Hypersonic Vehicles." Frontiers in Physics, Plasma Physics, Frontiers, 17 Jan. 2019, www.frontiersin.org/articles/10.3389/fphy.2019.00009/full.
13. Ahmed, Waqas *et al*. "Hypersonic Flow over Cone and Wedge in context of Hypersonic Vehicle Design using CFD". IBCAST 34. 2010.
14. Epifanov, V. M. "BOUNDARY LAYER." THERMOPEDIA, thermopedia.com/content/595/.

JOURNAL OF
EMERGING INVESTIGATORS

Article

Characterization of antibacterial properties of common spices

Youssef Gehad[1], Mark Springel[2]
[1]Medford High School, Medford, MA 02155
[2]Harvard University, Cambridge, MA 02138

SUMMARY

Antibiotics are used by physicians to treat bacterial infections. However, due to overuse, antibiotic-resistant strains of bacteria now threaten the efficacy of many antibiotic types. We tested whether 11 commonly used spices could inhibit growth of the gram-negative bacteria, *E. coli.* We tested cinnamon, clove, thyme, oregano, cumin, garlic, black pepper, rosemary, basil, and ginger, all of which have previously been suggested to have antibiotic properties. When these spices were diluted in Luria-Bertani (LB) agar, five spices (clove, cinnamon, garlic, sage, and thyme) inhibited growth completely. After a second round of experimentation using these five spices in liquid LB cultures, clove inhibited growth at a low concentration most effectively and four other spices (cinnamon, garlic, thyme, and sage) also effectively inhibited growth. The results of this study suggest that certain spices and herbs have antibacterial effects that can inhibit growth of *E. coli* and that these spices could show similarly promising activity towards other bacteria.

INTRODUCTION

Bacterial infection is resurging as one of the most dangerous challenges facing the medical establishment. Americans spend about 55 to 70 billion dollars per year on antibiotics, yet these antibiotics are becoming increasingly ineffective as illness-causing bacteria gain resistance to the prescribed drugs (1). This increasing resistance is making it harder for physicians to prescribe antibiotics that can treat bacterial infections. Additionally, some bacterial strains have started to become resistant not only to single antibiotics, but to multiple classes of antibiotics, making these infections very difficult to treat and dangerous to public health (2).

Antibiotic resistance occurs when bacteria evolve to prevent the deleterious effects of antibiotics on their growth and survival. There are two means by which bacteria resist the effects of antibiotics: an innate ability of the bacterial species to resist a given antibiotic (intrinsic resistance) or the acquisition of resistance through either random DNA mutation or horizontal transfer of resistance genes from another bacterial species (acquired resistance) (4). One commonly proposed hypothesis about the origin of innate resistance is that because some bacteria produce antibiotic compounds to compete with other bacterial species, these bacteria should have resistance to that antibiotic, so they do not kill themselves. These resistance genes may have first been transferred to pathogenic bacteria through horizontal transfer (3). However, resistance has now been propagated through random DNA mutation due to the pressures of the overuse of antibiotics (5).

For centuries before the discovery of antibiotics, humans relied on medicines made from herbs and spices. Furthermore, people used and passed down remedies which typically consisted of substances people ate or added to their food (6). With the rise of antibiotic resistance, there has been an increased interest in plants as an economical and renewable source of antimicrobials with rich chemical diversity (7).

Most bacteria that are resistant to one or more antibiotics are gram-negative, meaning that they have only a thin layer of peptidoglycan between their inner and outer cell membrane. Since most antibiotics have to pass through the outer membrane of bacteria, any change in the outer membrane by the gram-negative bacteria can create antibiotic resistance (4). Therefore, many research studies that test novel antibiotics use gram-negative bacteria, such as *E. coli*, as their model system. Unfortunately, antibiotic resistance is still a major problem for many gram-positive bacteria, including methicillin-resistant Staphylococcus aureus (MRSA).

We evaluated whether spices and herbs could be used as an effective method to suppress bacteria growth and in turn antibiotic resistant bacteria. We tested 11 spices and herbs: cinnamon, clove, thyme, oregano, cumin, garlic, black pepper, rosemary, basil, sage, and ginger. We used these 11 spices, because each of them was previously suggested to have antibacterial properties; as a result, we wondered which ones were most effective at suppressing growth of *E. coli* bacteria (8). While we hypothesized that all the spices would be capable of suppressing bacteria growth in a concentration-dependent manner, we found that only five spices are most effective at preventing bacteria growth in both Luria-Bertani (LB) agar and liquid cultures: clove, cinnamon, garlic, sage, and thyme. This result is a good sign that some spices and herbs could potentially be useful in combating the rise of antibiotic resistant bacteria.

RESULTS

To determine which spices might affect *E. coli* growth, we grinded each of the 11 spices to a fine consistency, mixed them into LB agar at 5% concentration, and made LB plates

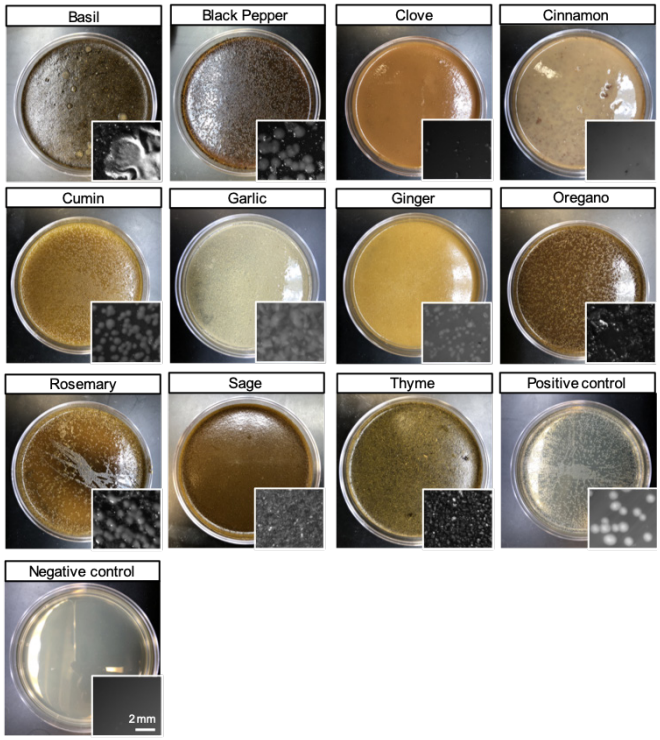

Figure 1: Images of *E. coli* grown with spice and herbs diluted in agar at 5% concentration. The images were acquired 18-24 hours after inoculation, following incubation of agar plates at 37°C.

using this agar. Equal amounts of *E. coli* were spread on each plate, three replicates of which were incubated overnight at 37°C. We chose to use a high, 5% concentration of spice in LB agar to ensure that spices with antibacterial properties would sufficiently inhibit bacterial growth as to be unambiguous. We chose a positive control of *E. coli* bacteria without any spice to ensure that the bacteria would grow normally. As a negative control, we conducted all steps of the experiment without adding *E. coli* to normal LB agar.

We found that four spices did not prevent bacterial growth: pepper, rosemary, cumin, and ginger (**Figure 1**). Interestingly, not all growth looked the same, and growth on the plates varied in both colony density and single colony size. Of these spices, ginger-treated plates had the lowest average colony density at 2.5 ± 0.4 (mean ± SEM) colonies per square millimeter (C/mm^2), while cumin, pepper, and rosemary exhibited 4.1 ± 0.2 C/mm^2, 2.7 ±0.7 C/mm^2, and 3.5 ± 0.5 C/mm^2, respectively (**Figure 2**). Plates mixed with 5% basil exhibited few if any E. coli colonies, and instead showed white 'bubble' growths that appeared fungal rather than bacterial. These colonies were irregular and not possible to quantify. The positive control plates exhibited a colony density of 2.5 ±0.7 C/mm^2, with a representative colony size of 0.86 mm^2 (**Figure 3**). Interestingly, agar plates treated with ginger and cumin both produced visually smaller colonies than positive control plates, with a representative colony measured at 0.15 mm^2 and 0.47 mm^2, respectively (**Figure 3**).

We found that the only plates that did not show any

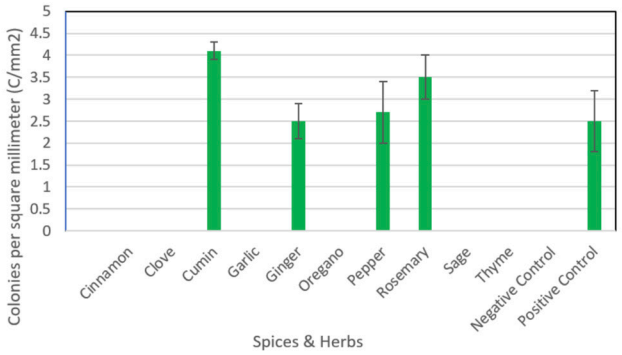

Figure 2: Mean *E. coli* colony density from cultures treated with spice and herb treatments, 18-24 hours after inoculation. Errors bars represent SEM, n=3 plates per treatment.

bacterial growth were clove, cinnamon, garlic, oregano, sage, and thyme, and the negative control. These plates lacked bacterial colonies entirely on all three replicate plates (**Figures 1 & 2**). We concluded that these five spices (cinnamon, garlic, thyme, sage, and clove) demonstrated the most potent antibacterial qualities.

Encouraged by the apparent potency of clove, cinnamon, garlic, oregano, sage, and thyme, we performed a follow-up experiment to determine whether these spices are effective at inhibiting bacterial growth at lower concentrations. The purpose of this second experiment was to find the minimum inhibitory concentration (MIC) of these spices, or the lowest amount of these spices needed to prevent bacteria growth. To find the MIC, we used a spectrophotometer to measure the optical density (OD) of *E. coli* LB cultures with each spice at increasing logarithmic concentrations of 0.01%, 0.1%, and 1%. We mixed the different concentrations of each spice in LB (with three replicates) and grew *E. coli* bacteria in the mixtures (**Figure 4A**). After incubation overnight, we centrifuged the bacteria into a pellet, resuspended it, and measured optical density using a spectrophotometer, which measures how much light of various wavelengths is transmitted through a liquid sample. By measuring light absorbance at 600nm, we estimated how many bacteria cells in colony-forming units (CFU) were present in the liquid, therefore gaining a measure of antibacterial activity for each spice.

We found that cinnamon, clove, sage, thyme, and garlic all inhibited bacteria growth when used at 1% concentration ($p < 0.05$, as assessed by Student's T-test), but not at 0.1% nor 0.01% (**Figure 4B**). Clove was the most effective at suppressing bacterial growth, and at 0.1% concentration also inhibited growth as compared to control (p=0.012, as assessed by Student's T-test), whereas garlic, sage, cinnamon, and thyme did not show statistically-significant decreases in OD600 at this concentration. Surprisingly, 1% oregano was ineffective at suppressing bacteria growth, and *E. coli* cultures with oregano exhibited higher bacteria concentrations than the positive control (0.393 vs. 0.283 OD600; p=0.024)

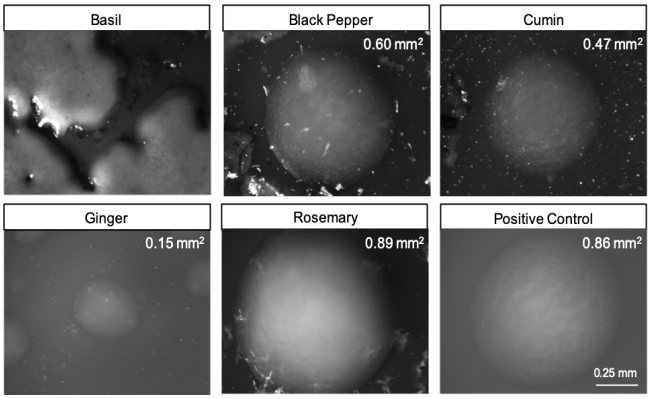

Figure 3: Representative images of individual colonies grown on agar plates with 5% spice treatment. Ginger and cumin treatments resulted in colonies smaller than those of control plates.

(**Figure 4**). The negative control demonstrated little to no bacteria growth, as assessed by visual clarity and OD600 measurements (**Figure 4**).

DISCUSSION

We hypothesized that all spices would be capable of suppressing *E. coli* growth, although in a concentration-dependent manner. Instead, we found that only five spices were effective in suppressing growth, and with the concentrations that we tested, growth suppression occurred in a binary fashion (growth or no growth). However, the fact that some spices and herbs were able to inhibit most bacteria growth at 1% concentrations suggests that other approaches to using spices – such as using fresh spices or extracting the active ingredients – might help in finding solutions to the antibiotic-resistance crisis. Further research that identifies these active ingredients might provide the basis for new antibiotic strategies to be integrated into the medical field. These new antibiotics could even be used in formulations with current antibiotics to overcome antibiotic-resistant bacteria strains.

Metabolites of the spice plants are likely the reason behind their antimicrobial properties (9). Typically, metabolites can inhibit bacterial growth by binding to bacterial proteins or by reducing the pH of the bacteria, which kills them (7). Additionally, metabolites possess antimicrobial mechanisms that can damage microbial membranes, impair cellular metabolism, and lower microbial toxin production (9). Interestingly, plant extracts have antimicrobial activity with potential to help prevent cell wall construction, impede microbial DNA replication, and constrain energy production (9). Modern medicine could harness these metabolites to prevent antibacterial resistance (9).

Previous research efforts have focused on combating the rise of antibiotic resistance or finding an alternative to antibiotics altogether. For example, researchers have studied the use of bacteriophages to combat the antibiotic-resistant bacteria, which can target and kill a specific bacteria strain while not harming eukaryotic cells of the host animal (4).

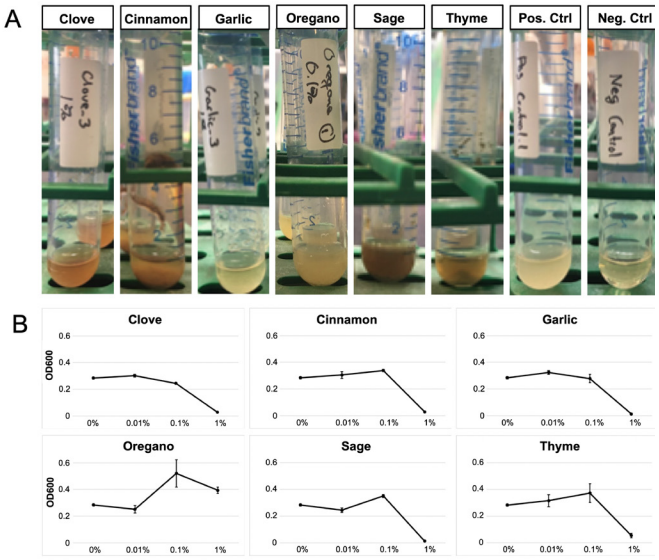

Figure 4: Dose response curves of *E. coli* grown in LB with various concentrations of spices. A) Images from *E. coli* cultures treated with 1% of six spices that demonstrated antibacterial activity in the previous experiment. B) OD600 values of *E. coli* cultures treated with 0%, 0.01%, 0.1%, and 1% of each spice. Errors bars represent SEM, n=3 cultures per treatment.

Another possible counter to antibiotic resistance is the use of antibiotic adjuvants, which on their own have no antibiotic use but are able to enhance those of antibiotics when combined as part of a drug therapy (4).

Several of our results were surprising to us, including the atypical growths seen in our basil LB agar cultures, and the lack of more nuanced spice dose-dependence of our LB culture dilutions. One explanation for the atypical basil culture growths is that the dried basil was contaminated with some other bacteria or fungus prior to dilution in the agar plates. To assess whether this result was consistent with basil from a different source, we might have tested dried basil from another store brand, or even tested fresh basil. Although basil was the only spice that demonstrated these atypical growths, it is possible that contamination from other microbes might also have influenced our results. Future studies should explore autoclaving the spices, or otherwise extracting active compounds from the spices before testing. For the second experiment with the liquid LB cultures, an alternative method might have been to wait 24 hours instead of just 18 hours of incubation, as it is possible that the spices slowed bacteria growth, rather than killing bacteria entirely. This may have allowed us more accurate distinction between spices at 1% concentration because some OD600 values were too low to determine if bacteria were present. If we were to repeat this experiment, we would also test intermediate spice concentrations between 0.1% and 1%, in hopes of establishing a more detailed dose-dependence curve.

Future studies should test different spices or herbs,

beyond what we tested, to widen the knowledge of effective plant metabolites against bacteria. Follow-up experiments should also attempt different methods of extraction to concentrate the active compounds and test them independently against *E. coli.* When a compound is identified and isolated, we could use a modern antibiotic in the liquid culture and compare how effective it was at low concentration to active compounds identified from the spices and herbs. Another experiment would be to use a different bacteria strain and see if the most effective spices and herbs are also potent against other gram-negative bacteria, or even gram-positive bacterial species. Lastly, we could test if *E. coli* might eventually develop resistance to a specific spice or herb, and how long would it take to do so.

MATERIALS AND METHODS

Preperation of herbs and spices

First, we ground the dried spices and herbs (purchased from Market Basket) using a coffee grinder until they were fine in quality. For those spices already ground to a fine consistency in their packaging, such as cinnamon and black pepper, we did nothing. Then, we put all spices in labeled falcon tubes to be stored as we conducted the experiment.

Preperation of LB agar plates

To make LB agar plates with 5% spice dilution, we melted 1000 mL of an LB agar stock (FisherSci) in a microwave until boiling. While we waited for the melted LB agar to cool, we labeled 39 petri dishes with the name of spice/positive control/negative control and trial number. We then poured 60 mL of the LB agar into secondary container with 3 g of the spice (5% concentration), mixed the LB agar - spice suspension, and poured 20 mL into each of three replicate 10 cm petri dishes. The control plates were poured with just the LB media agar by itself. We let the LB agar in petri dishes solidify overnight.

Inoculation and incubation of LB agar plates

To grow a stock of *E. coli* cells, DH5-alpha *E. coli* (New England Biolabs) cells were thawed at room temperature, and 5 µL were pipetted into 2 mL of LB media in a test tube. The *E. coli* were then incubated overnight at 37°C and 220 revolutions per minute. The next day, we pipetted 2 µL *E. coli* into a NanoDrop spectrophotometer to measure number of bacteria. Then, we diluted the *E. coli* cultures with LB media to produce an estimated 2.0 x 104 CFU per mL concentration. We pipetted 20 µL of the *E. coli* solution into the center of each plate, and spread the *E. coli* evenly over each plate using sterile glass plating beads. Last, we incubated the plates overnight at 37°C and checked their growth the next day.

Colony counting and analysis

For the analysis, we took representative images spanning 308.5 mm^2 of the plates under a dissection microscope (Zeiss AxioZoom). Using the Fiji application, we manually counted the number of bacteria on the images of the spices that had visible growth of bacteria on each of their three plates. Additionally, we took a picture of one individual representative colony from each spice treatment. We used the Fiji application to measure the area of the representative colony.

Inoculation and incubation of LB liquid cultures

First, we measured out 10 mL of LB into 39 Falcon tubes (3 tubes for each condition). Next, we measured out 100 mg, 10 mg, and 1 mg for each spice (1%, 0.1%, and 0.01%). We mixed the spice with the LB media and vortexed it for 1 minute. Then, we pipetted 2 mL from each Falcon tube into culture tubes with 2000 CFU (1000 CFU/µl). Finally, we incubated the tubes for 18 hours at 37°C and 220 revolutions per minute.

Optical density readings and analysis

The following day, we centrifuged the tubes at 10 xG for 1 minute to pellet small spice particulate at the bottom of each tube, and collected 1 mL of the supernatant into new Eppendorf tubes with no spice particles. Then, we centrifuged the supernatant at 16000 xG for 5 minutes to pellet all *E. coli* bacteria. We poured out the remaining liquid in the Eppendorf tubes, leaving only the *E. coli* pellet, and resuspended the pellet in 1 mL Phosphate Buffered Saline (PBS) .

Subsequently, we measured 2 µL of the resuspension onto a NanoDrop spectrophotometer and recorded the OD600 value for each sample. Then, we used an online calculator to estimate CFU from each sample. After that, we found the mean and standard error for three trials of each treatment condition. Data analysis and statistics were performed on the raw OD600 values using Microsoft Excel.

Received: May 11, 2020
Accepted: Jun 15, 2020
Published: October 03, 2020

REFERENCES

1. "The NCES Fast Facts Tool Provides Quick Answers to Many Education Questions (National Center for Education Statistics)." National Center for Education Statistics (NCES) Home Page, a Part of the U.S. Department of Education, nces.ed.gov/fastfacts/display.asp?id=372.
2. Nikaido, Hiroshi. "Multidrug Resistance in Bacteria." Annual Review of Biochemistry, vol. 78, 2009, pp. 119–146. doi:10.1146/annurev.biochem.78.082907.145923.
3. Ogawara, Hiroshi. "Comparison of Antibiotic Resistance Mechanisms in Antibiotic-Producing and Pathogenic Bacteria." Molecules, vol. 24, 2009, pp. 1-55. doi:10.3390/molecules24193430
4. Breijyeh, Zeinab et al. "Resistance of Gram-Negative Bacteria to Current Antibacterial Agents and Approaches to Resolve It." Molecules, vol. 25, 2020, pp. 1-23. doi:10.3390/molecules25061340
5. Harnisz, Monika et al. "tet genes as indicators of changes in the water environment: relationships between culture-dependent and culture-independent approaches." Sci

Total Environ, vol. 505, 2015, pp. 704-711. doi:10.1016/j.scitotenv.2014.10.048
6. McCormick Science Institute. "History of Spices." McCormick Science Institute, www.mccormickscienceinstitute.com/resources/history-of-spices.
7. Ruddaraju, Lakshmi Kalyani et al. "A review on anti-bacterials to combat resistance: From ancient era of plants and metals to present and future perspectives of green nano technological combinations." Asian J Pharm Sci, vol. 15, 2020, pp. 42-59. doi:10.1016/j.ajps.2019.03.002
8. Liu, Qing et al. "Antibacterial and Antifungal Activities of Spices." Int J Mol Sci, vol. 18, 2017, pp. 1-62. doi:10.3390/ijms18061283
9. Mickymaray, Suresh. "Efficacy and Mechanism of Traditional Medicinal Plants and Bioactive Compounds against Clinically Important Pathogens." Antibiotics, vol. 8, 2019, pp. 1-57. doi:10.3390/antibiotics8040257
10. "E. Coli Cell Culture Concentration from OD600 Calculator." Agilent Genomics: Tools - Bio Calculators, www.chem.agilent.com/store/biocalculators/calcODBacterial.jsp.

ACKNOWLEDGEMENTS

I would like to thank Harvard Medical School for letting us use lab equipment for both of our experiments. Also, I would like to thank the JEI Mini-PhD program for giving me a chance to develop an independent experiment and pairing me with a thoughtful, creative, and intelligent mentor to guide me in this process.

JOURNAL OF
EMERGING INVESTIGATORS

Article

An Analysis of the Density and Patterns of the Solutions of Diophantine Equations of the Third Power

Sidak Singh Grewal[1] & Kanwarpreet Grewal[2]
[1] Lotus Valley International School, Noida, India
[2] Express Greens, Sector 44, Noida, India

SUMMARY

According to Fermat's last theorem, $x^n + y^n = z^n$ has no solutions if $n > 2$. We modified Fermat's equation into the Diophantine equations $a^3 + b^3 + c^3 = d^3$, $a^3 + b^3 + c^3 = d^2$ and $a^3 + b^3 + c^3 = d^4$ and found their solutions. We analyzed how the density of solutions varied as the numbers got bigger. Ramanujan had devised a formula to find numbers satisfying $a^3 + b^3 + c^3 = d^3$, we compared the density of solutions with those obtained by his formula. We also found perfect cubes, squares or fourth power solutions that could be expressed in different ways as a sum of three cubes. We called them perfect power taxicab numbers. Our hypothesis was that there are many solutions for our equations, and as the inputs become bigger, their density will increase linearly with minor fluctuations. We thought that most perfect power taxicab numbers would have a frequency (number of ways it can be expressed as a sum of three cubes) of two and the maximum frequency would be around 10. We hypothesized that Ramanujan's formula would give around half of the solutions, and the density of solutions will increase as the numbers become large. We concluded that the density distribution of two equations increases as the numbers become bigger. However, the third equation had a stagnant density. Ramanujan's formula found many numbers at the start but was unable to reach a high density. One perfect cube taxicab number had a frequency of 42 whereas the majority had a frequency of 2 or 3.

INTRODUCTION

Since the dawn of civilization, we have been obsessed with finding patterns in the world around us. Our ancestors created mathematics to solve daily problems but were fascinated to find never-ending patterns in numbers. The entire branch of number theory is focused on finding interesting patterns in integers.

Diophantine equations are polynomial equations with more than one variable in which we seek only integral solutions. An exponential Diophantine equation is an equation with exponents on the variables (1). Pythagoras's theorem states that the lengths of the sides of a right-angle triangle satisfy the Diophantine equation $b^2 + p^2 = h^2$. We know that this equation has infinite solutions (2).

Pierre de Fermat was one of the greatest number theorists of the 17th century. On the margin of a copy of the ancient Greek book Arithmetica, he conjectured that the Diophantine equation $x^n + y^n = z^n$ has integer solutions only if n is less than or equal to 2 (3). This statement came to be known as Fermat's last theorem. For example, Fermat's last theorem states that the Diophantine equations $x^3 + y^3 = z^3$ or $x^4 + y^4 = z^4$ do not have any integral solutions. This conjecture turned out to be extremely difficult to prove and this proof eluded mathematicians for centuries. This conjecture was finally proven in 1995 in a 129-page paper by Sir Andrew Wiles (4). Ramanujan, the famous Indian mathematician, contributed to several fields of mathematics including number theory, continued fractions, mathematical analysis and infinite series (5,6). His works continue to be studied and analyzed by mathematicians even a century after he died. Many mathematicians including Ramanujan devised formulas to find numbers satisfying $a^3 + b^3 + c^3 = d^3$ (7,8). There is no known formula that can find all the solutions of the above equation (9).

In the famous exchange when G. H. Hardy came to visit Ramanujan in the hospital, Hardy said that the cab he came to meet him in was numbered 1729. Hardy remarked it was an extremely dull number (10). Ramanujan however, answered that it was a very interesting number. Ramanujan said that this was the smallest integer which could be expressed as the sum of two cubes in two different ways:

$$1729 = 1^3 + 12^3$$
$$1729 = 9^3 + 10^3$$

Numbers of this kind came to be known as taxicab numbers or Hardy-Ramanujan numbers (11).

In this paper we have varied the equation in Fermat's last theorem into Diophantine equations of the form $a^3 + b^3 + c^3 = d^3$, $a^3 + b^3 + c^3 = d^2$ and $a^3 + b^3 + c^3 = d^4$ and found integral solutions for them keeping a, b and c to less than 10,000. Using a computer program that we wrote, we found many sets of integers satisfying the above three equations by brute force – putting in each number one by one and only keeping the sets that satisfy the equations. After we found all the solutions under 10,000, we analyzed the density of solutions and how it varied as the numbers got bigger.

We wanted to find out how many solutions Ramanujan's formula could find, and we wanted to compare the density of solutions got by his formula to our own calculated density.

To accomplish this, we made a computer program to find the solutions of $a^3 + b^3 + c^3 = d^3$ using his formula and compared the density of solutions of those obtained by his formula with those we found.

In this paper, we generalized Hardy-Ramanujan's taxicab numbers to a perfect power able to express itself as a sum of three cubes in different ways. For example, the number 1,000,000 is a perfect cube and can be expressed in two ways as a sum of three cubes:

$$1{,}000{,}000 = 100^3 = 85^3 + 70^3 + 35^3$$
$$1{,}000{,}000 = 100^3 = 88^3 + 68^3 + 16^3$$

So, we say that 1,000,000 is a perfect cube taxicab number and has a frequency of 2.

We found perfect square, perfect cube and perfect fourth power taxicab numbers keeping *a*, *b* and *c* to less than 10,000.

Our perfect power taxicab numbers are different from Ramanujan's taxicab numbers in the way that our numbers are the sum of three and not two cubes and the result itself is a perfect cube, perfect square or perfect fourth power rather than any integer as in Ramanujan's case. We found many perfect cubes, square and fourth power taxicab numbers according to the equation. We analyzed and compared the frequency of perfect power taxicab numbers according to the number of times they could be expressed in different ways.

Our hypothesis was that there are many solutions for our equations, and as the inputs become bigger, their density will increase linearly with minor fluctuations. We thought that most perfect power taxicab numbers would have a frequency of two and the maximum frequency would be around ten. We hypothesized that Ramanujan's formula would give around half of the solutions, and the density of solutions will increase as the numbers become large.

RESULTS

We analyzed the solutions of the three equations $a^3 + b^3 + c^3 = d^3$, $a^3 + b^3 + c^3 = d^2$ and $a^3 + b^3 + c^3 = d^4$ and compared their densities.

From our data we found interesting patterns in all the three equations. The number of solutions of the equation $a^3 + b^3 + c^3 = d^3$ in which all *a*, *b* and c are less than or equal to 10,000 is 65,085. Throughout our analysis, we assumed that $a \geq b$ and c. Therefore, we used a (the largest of the three inputs) as the basis of our density analysis. We found that the number of solutions of $a^3 + b^3 + c^3 = d^2$ for *a*, *b* and c less than or equal to 10,000 was 266,517 and that of the equation $a^3 + b^3 + c^3 = d^4$ for *a*, *b*, and *c* less than or equal to 10,000 was just 987.

Density of solutions for $a^3 + b^3 + c^3 = d^3$ (Cubic equation)

The total number of solutions for *a*, *b*, and c below or equal to 10,000 were 65,085. The first solution was

$$5^3 + 4^3 + 3^3 = 6^3 = 216,$$

and the last solution in the set of numbers that we considered was

$$9999^3 + 9696^3 + 9494^3 = 14039^3 = 2{,}766{,}995{,}941{,}319.$$

We used our data to analyze how the density of solutions to this equation varies as the input numbers (*a*, *b*, and *c*) become bigger. To visualize the density of solutions, we looked at the solutions in two different ways. We considered the number of solutions for every 100 multiples of the largest of *a*, *b*, and *c* in our equation. In our analysis we always took $a \geq b$ and c. For example, between $a = 0$ and $a = 100$, there were 98 solutions. Between $a = 100$ to $a = 200$, there were 183 solutions, and so on (**Figure 1a**). We also analyzed how the number of solutions per number of combinations of *a*, *b* and *c* varies per 100 multiples of *a* (**Figure 1b**).

We could see that the number of solutions per hundred values of a (density of solutions) rose quite steeply at first and then the rise became more gradual (**Figure 1a**). As we approached $a = 10{,}000$, the density reached more than eight hundred solutions for every 100 values of *a*. There were some minor fluctuations, but the overall trend was that of a rapid increase in the beginning and a slower increase later. The maximum number of solutions was found between $a = 9000$ and $a = 9100$. This range has 902 solutions. The minimum number of solutions was between $a = 0$ and $a = 100$: 98 solutions (**Figure 1a**).

As *a*, *b* and *c* increased the total number of solutions per combinations of *a*, *b* and *c* decreased rapidly (**Figure 1b**). In a sense total number of solutions per combinations

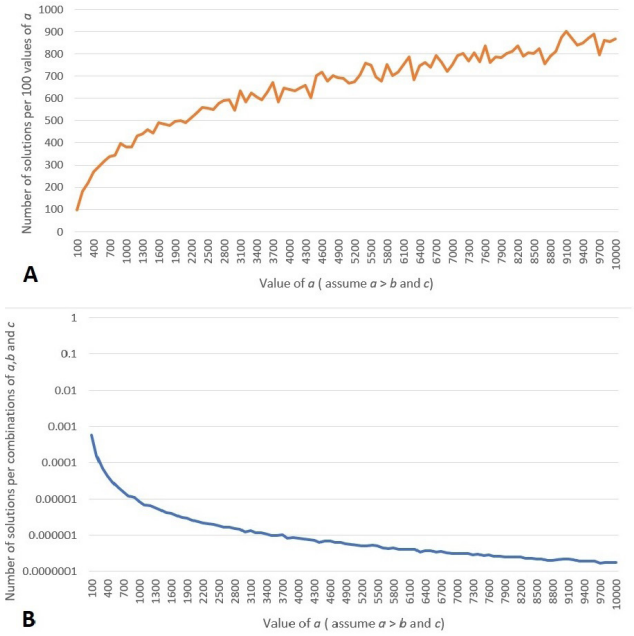

Figure 1: Solution density for the cubic Diophantine equation. A) Number of solutions per hundred numbers of *a* for the equation $a^3 + b^3 + c^3 = d^3$. The maximum number of solutions is found between $a = 9000$ and $a = 9100$. This range has 902 solutions. The minimum number of solutions is between $a = 0$ and $a = 100$: 98 solutions. B) The number of solutions per number of combinations of *a*, *b* and *c* vs values of *a* with bin size of 100 ($a \geq b$ and *c*).

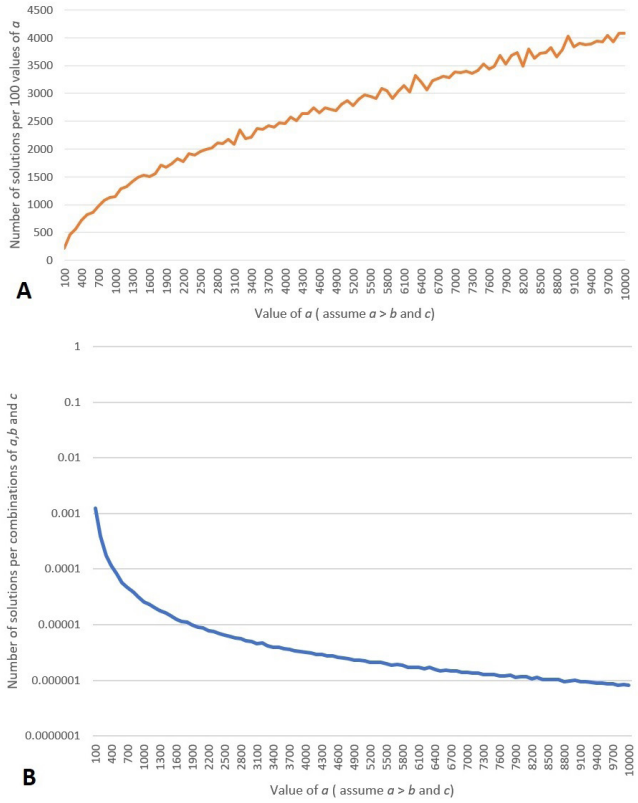

Figure 2: Solution density for the square Diophantine equation. A) Number of solutions per hundred numbers of *a* for the equation $a^3 + b^3 + c^3 = d^2$. The minimum density in the range of numbers that we considered was between $a = 0$ to $a = 100$, which was 214. The maximum density was between $a = 9{,}800$ to $a = 9{,}900$, which was 4,078. B) The number of solutions per number of combinations of *a*, *b* and *c* vs values of *a* with bin size of 100.

of the inputs (**Figure 1b**) is a better measure of the density compared to looking at just the largest input (**Figure 1a**). The latter only focuses on the largest variable *a*. It does not take into account the fact that as a increases, the number of *b* and *c* values that can be tried also increase. However, **Figure 1b** normalizes the density by finding the total number of solutions in any given range of a divided by number of combinations of *a*, *b* and *c* that will be tried in that range of *a*.

Density of solutions for $a^3 + b^3 + c^3 = d^2$ (Square or quadratic equation)

The total number of solutions below ten thousand values of a was 266,517. The first solution was

$$3^3 + 2^3 + 1^3 = 6^2 = 36,$$

and the last solution in our data was

$$9999^3 + 9945^3 + 6390^3 = 1498068^2 = 2{,}244{,}207{,}732{,}624.$$

We used our data to analyze how the density of solutions to this equation varied as the input numbers (*a*, *b* and *c*) become bigger. We considered the number of solutions for every 100 multiples of the largest of *a*, *b* and *c* in our equation (**Figure 2a**). In our analysis we always took $a \geq b$ and *c*. The minimum density in the range of numbers that we considered was between $a = 0$ to $a = 100$, which was 214. The maximum density was between $a = 9{,}800$ to $a = 9{,}900$, which was 4,078. From the plot we concluded that as the value of a increased, the density of solutions also steadily increased.

We also analyzed how the number of solutions per combinations of *a*, *b* and *c* varied per 100 multiples of *a* (**Figure 2b**). The number of solutions per combinations of *a*, *b* and *c* dropped quite rapidly at first and then flattened out.

Density of solutions for $a^3 + b^3 + c^3 = d^4$ (4^{th} power equation or quartic equation)

The total number of solutions of this equation within the first 10,000 values of *a*, *b* and *c* is only 987. The first solution was

$$3^3 + 3^3 + 3^3 = 3^4 = 81$$

and the last solution within the numbers that we considered was

$$9986^3 + 4544^3 + 3006^3 = 1028^4 = 1{,}116{,}792{,}422{,}656$$

The total number of solutions decreased from the quadratic to cubic to quartic (266517 vs 65085 to 987), and that was expected because the total count of numbers whose square is less than 3×10^{12}(*a*, 10^{12} values of *b* and 10^{12} values of *c*) is much more than the total number of cubes in the range which, in turn, is much greater than the total number of powers of four in the range.

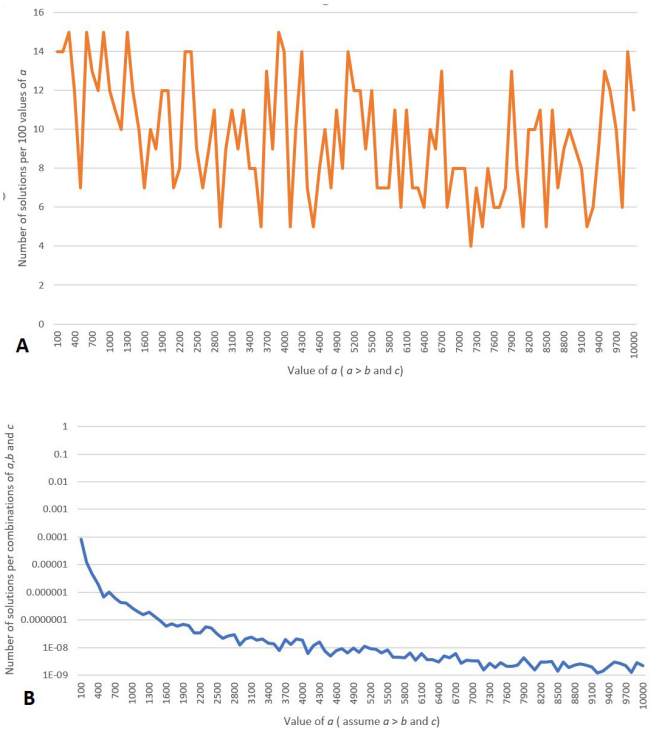

Figure 3: Solution density for the quartic Diophantine equation. A) Number of solutions per hundred numbers of *a* for the equation $a^3 + b^3 + c^3 = d^4$. The maximum solution density is 15, which is achieved by five ranges of *a*. The minimum frequency density was between $a = 7100$ and $a = 7200$, which is four. This equation's solution density shows a lot of local fluctuations but has a constant "flat" trend. B). The number of solutions per number of combinations of *a*, *b* and *c* vs values of *a* with bin size of 100.

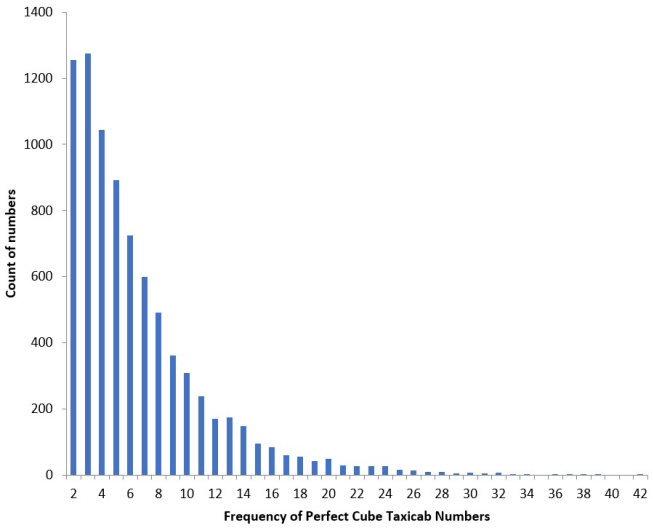

Figure 4: Histogram which compares the count of numbers with different frequencies for the equation $a^3 + b^3 + c^3 = d^3$. The count of numbers with frequency of 3 is greater than count of numbers with frequency of 2. The maximum frequency is 42 for perfect cube taxicab numbers

We analyzed the solution density of this equation for the first 10,000 values of *a* (assume $a \geq b$ and *c*) (**Figure 3a**). We also considered how the number of solutions per number of combinations of *a*, *b* and *c* varied per 100 multiples of *a* (**Figure 3b**).

We inferred (**Figure 3a**) that this equation's solution density was unlike that of the previous two equations (the square and the cubic). Here the maximum solution density was 15, which was achieved by five ranges of a. The minimum density was between $a = 7100$ and $a = 7200$, which was four. This equation's solution density showed a lot of local fluctuations but had a constant "flat" trend. The first range $a = 0$ to $a = 100$ had the highest density of 15, and this was repeated 4 times. This curve was constrained within the range 4-15 with large fluctuations. In the square and cubic results (Figures 1a and 2a), the value of the density went up to thousands of numbers while the highest density for the quartic equation only went up to 15.

When we analyzed how the number of solutions per number of combinations of *a*, *b* and *c* varied per 100 multiples of *a* we saw that this shows a different character than the corresponding solutions of the square and cubic equation (**Figure 3b**). The fall was faster and the flattening happened later compared to the square and the cubic equation. The density also showed significant local variations.

A very effective way to compare the three equations was to fit some simple polynomial curves on the data and compare these curves. We found that the following three curves fit into the data of number of solutions per 100 values of *a* (Figures 1a, 2a and 3a):

Cubic equation (**Figure 1a**):
$-6.312\times10^{-6}\, x^2 + 0.1199x + 258.9$

Square equation (**Figure 2a**):
$-2.145\times10^{-5}\, x^2 + 0.5471x + 627.6$
4th power equation (**Figure 3a**):
$1.132\times10^{-7}\, x^2 - 0.001511x + 13.35$

When we compare these three polynomials by plotting them and by algebraic analysis, we saw that the fourth power polynomial was almost flat. The third power curve had a more gradual slope than the 2nd power curve and did not rise as high.

The following curves fit the data of the number of solutions per combinations of *a*, *b* and *c* plotted against values of a with bin size of 100 (Figures 1b, 2b and 3b)

Cubic equation (**Figure 1b**):
$3.365\times10^{-8}\, x^2 - 0.0005524\, x - 4.393$
Square equation (**Figure 2b**):
$3.084\times10^{-8}\, x^2 - 0.0005014\, x - 3.965$
4th power equation (**Figure 3b**):
$4.69\times10^{-8}\, x^2 - 0.0007486\, x - 5.683$

When we compared these polynomials we saw that the fourth power curve fell much more rapidly compared to the other curves and reached a much lower point than the other two curves. The flattening of this curve happened later compared to the cubic equation and square equation curves. The cubic curve and the square equation curve had similar slopes and shape but the cubic curve was lower than the square equation curve but still much higher than the 4th power curve.

Perfect Power Taxicab numbers

We defined perfect power taxicab numbers as perfect cubes, squares or fourth powers satisfying the Diophantine equations outlined above, which can be expressed as a sum of three cubes in different ways. Surprisingly we found many perfect power taxicab numbers for all the three Diophantine equations that we considered.

We found perfect cube taxicab numbers i.e. multiple *a*, *b* and *c* for the same *d* satisfying $a^3 + b^3 + c^3 = d^3$ and compared the count of numbers for each frequency (**Figure 4**). We

$970299000000 = 9900^3$		
$= 7590^3 + 6490^3 + 6380^3$	$= 7625^3 + 6775^3 + 6000^3$	$= 7704^3 + 6796^3 + 5840^3$
$= 7911^3 + 7713^3 + 2538^3$	$= 8019^3 + 6822^3 + 5157^3$	$= 8040^3 + 7512^3 + 2988^3$
$= 8145^3 + 7240^3 + 3695^3$	$= 8205^3 + 6030^3 + 5835^3$	$= 8208^3 + 6156^3 + 5688^3$
$= 8240^3 + 6540^3 + 5080^3$	$= 8250^3 + 6600^3 + 4950^3$	$= 8310^3 + 6810^3 + 4320^3$
$= 8355^3 + 6885^3 + 3930^3$	$= 8415^3 + 6930^3 + 3465^3$	$= 8485^3 + 6939^3 + 2936^3$
$= 8487^3 + 6504^3 + 4377^3$	$= 8500^3 + 6750^3 + 3650^3$	$= 8586^3 + 6936^3 + 1542^3$
$= 8614^3 + 6848^3 + 2154^3$	$= 8640^3 + 6750^3 + 2610^3$	$= 8646^3 + 6858^3 + 1128^3$
$= 8712^3 + 5676^3 + 5016^3$	$= 8712^3 + 6732^3 + 1584^3$	$= 8766^3 + 6108^3 + 4098^3$
$= 8800^3 + 6600^3 + 1100^3$	$= 8982^3 + 6186^3 + 2076^3$	$= 9036^3 + 6108^3 + 1668^3$
$= 9108^3 + 4818^3 + 4686^3$	$= 9120^3 + 5170^3 + 4190^3$	$= 9225^3 + 5175^3 + 3600^3$
$= 9225^3 + 5700^3 + 375^3$	$= 9231^3 + 4593^3 + 4428^3$	$= 9384^3 + 4491^3 + 3765^3$
$= 9416^3 + 4785^3 + 2959^3$	$= 9475^3 + 4925^3 + 600^3$	$= 9570^3 + 4180^3 + 2750^3$
$= 9609^3 + 4054^3 + 2543^3$	$= 9636^3 + 4180^3 + 1364^3$	$= 9675^3 + 3725^3 + 2350^3$
$= 9719^3 + 3313^3 + 2514^3$	$= 9850^3 + 2300^3 + 1350^3$	$= 9890^3 + 1360^3 + 750^3$

Figure 5: 42 ways in which a perfect cube could be expressed as a sum of three different cubes.

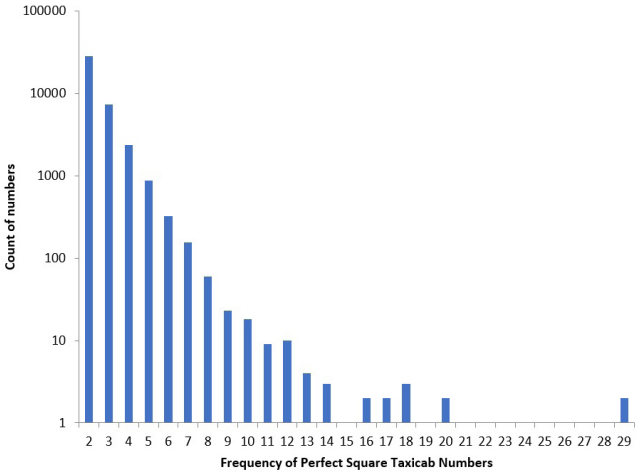

Figure 6: Histogram which compares the count of numbers with different frequencies for the equation $a^3 + b^3 + c^3 = d^2$. The maximum frequency is limited to 29 for perfect square taxicab numbers.

define the frequency of a taxicab number for the equations we considered as the number of different ways a perfect cube, square or fourth power taxicab number is expressed as a sum of three cubes. For example:

$$1{,}000{,}000 = 100^3 = 85^3 + 70^3 + 35^3$$
$$1{,}000{,}000 = 100^3 = 88^3 + 68^3 + 16^3$$

Therefore, we say that 1,000,000 is a perfect cube taxicab number and has a frequency of 2. There are 1256 taxicab numbers which have a frequency of 2 for the equation $a^3 + b^3 + c^3 = d^3$. Thus, we say that the count of numbers that have the frequency of 2 is 1256.

Our calculations allowed us to find all perfect cube taxicab numbers up to $10{,}000^3$ ($d^3 \leq 10{,}000^3$). It may seem that our analysis should have found all perfect cube taxicab numbers up to $3 \times (10{,}000)^3$ since a ranges from 1 to 10,000 and b and c range from 0 to a, but that is not the case. There will be taxicab numbers less than $3 \times (10{,}000)^3$ but greater than $10{,}000^3$ which have $a > 10{,}000$ but b and $c < 10{,}000$. On the other hand, we can prove that all taxicab numbers up to $10{,}000^3$ have been covered in our analysis. This can be seen from the fact that the maximum value of a is 10,000 and the minimum value of b and c are 0. In this limiting case the equation becomes $10{,}000^3 + 0^3 + 0^3 = 10{,}000^3$. Thus, our analysis of a, b and c numbers where each of them is less than or equal to 10,000 does cover all perfect cube taxicab numbers up to $10{,}000^3$ but does not cover all perfect cube taxicab numbers $> 10{,}000^3$ as some of them could have $a > 10{,}000$ but b and $c < 10{,}000$. Therefore, although we had found some frequencies of taxicab numbers $d^3 > 10^{12}$ in our data, we did not consider it in our analysis as this data is incomplete for the above-mentioned reasons.

In total, there were 8276 different perfect cube taxicab numbers up to $a = 10{,}000$. Most of these numbers had a frequency of 2 and 3 but we could find many numbers whose frequency in this data went beyond twenty (**Figure 4**). The maximum frequency for a perfect cube taxicab number was 42. So, there were 42 ways in which a perfect cube could be expressed as a sum of three different cubes (**Figure 5**). We had expected that the count of numbers with frequency of 2 will dominate but we were surprised to see that the count of numbers with frequency of 3 was greater than the count of numbers with frequency of 2.

We then compared the count of numbers with different frequencies for the equation $a^3 + b^3 + c^3 = d^2$ (**Figure 6**). These are perfect square taxicab numbers i.e. multiple a, b and c for same d satisfying the equation. As described for the cubic Diophantine equation above, our data consists of all taxicab numbers such that $d^2 \leq 10{,}000^3$ (i.e. $d < 10^6$).

We inferred that although the number of perfect square taxicab numbers (39,153) is far greater than that of the perfect cube taxicab numbers (8276), the frequency is limited to a certain range (**Figure 4**, **Figure 6**). The maximum frequency is limited to 29 for perfect square taxicab numbers whereas the perfect cube taxicab numbers peaked at a frequency of 42. Another difference from the cubic case was that for the square equation, the count of numbers with frequency of 2 far exceeds the count with the frequency of 3.

We analyzed the frequency of perfect fourth power taxicab numbers, i.e. multiple a, b and c for same d satisfying $a^3 + b^3 + c^3 = d^4$ (**Figure 7**). We compared the count of numbers with different frequencies. As described in the cubic Diophantine equation above, we found all taxicab numbers

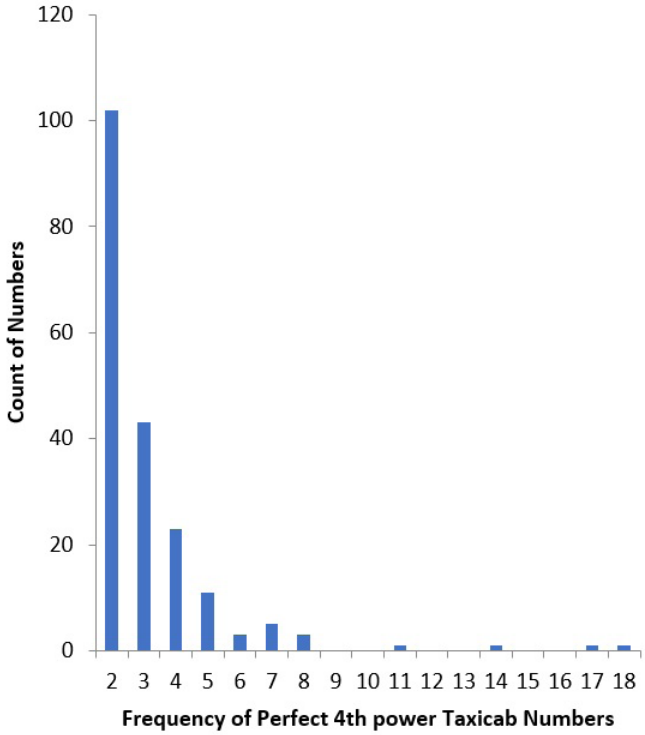

Figure 7: Histogram which compares the count of numbers with different frequencies for the equation $a^3 + b^3 + c^3 = d^4$. The maximum frequency of perfect fourth power taxicab numbers in the first 10,000 numbers was 18.

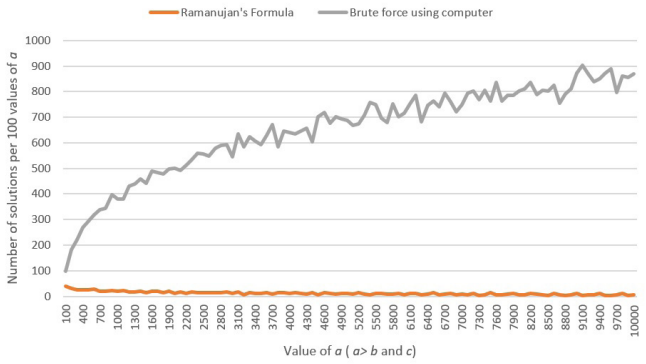

Figure 8: Comparison of the density of solutions of $a^3 + b^3 + c^3 = d^3$ got by Ramanujan's formula and our computer program (brute force method). Ramanujan's formula could only find a fraction of the solutions that we found by brute force, and as the numbers grew larger, the gap became bigger.

such that $d^4 = 10{,}000^3$ (i.e. $d = 10^3$).

We could see that there are very few fourth power taxicab numbers in the first ten thousand numbers of *a* (assuming $a \geq b, c$) (**Figure 7**). However, we saw that out of 987 solutions of this equation in the range of numbers we considered, there were 194 different perfect fourth power taxicab numbers. Therefore, the ratio of taxicab numbers to total solutions in the 4th power equation is much greater than the ratio in the square equation and the cubic equation. The maximum frequency of perfect fourth power taxicab numbers in the first 10,000 numbers was 18.

Ramanujan's formula for Diophantine equations of the third power

Ramanujan proposed a formula to find solutions of $a^3 + b^3 + c^3 = d^3$ (7,8) which is as follows:

$$(3x^2 + 5xy - 5y^2)^3 + (4x^2 - 4xy + 6y^2)^3 + (5x^2 - 5xy - 3y^2)^3 = (6x^2 - 4xy + 4y^2)^3$$

Here $(3x^2 + 5xy - 5y^2)$, $((4x^2 - 4xy + 6y^2)$ and $(5x^2 - 5xy - 3y^2)$ are our *a*, *b* and *c*, where the largest number for a specific *x*, *y* is *a*, second largest number is *b* and the smallest is *c*. To generate the solutions obtained by the above equation, we inserted values of *x* and *y* such that *a* would range from 0 to 10,000. Here both *x* and *y* can be negative but we selected only those solutions for which *a*, *b*, *c* and *d* were positive. We plotted the density of solutions of $a^3 + b^3 + c^3 = d^3$ generated by Ramanujan's formula and compared this data with our own data for a less than or equal to 10,000 (**Figure 8**). For this analysis we considered the number of solutions per hundred numbers of *a* and compared the density obtained from Ramanujan's formula to the density obtained by our program. A complete density analysis would also require us to divide the number of solutions in a given interval by the number of combinations of *a*, *b* and *c*. However, our main aim in this analysis was to compare Ramanujan's formula with the number of solutions found by brute force and that comparison would be equally relevant if we were just to consider the number of solutions per 100 values of *a*.

When *a* was small, the gap between the density of solutions found by Ramanujan's formula and our program was about 3 times: 38 vs 98 in the first 100. However, the density of solutions in the two cases quickly diverged. With the increase in *a*, the density of solutions of Ramanujan's formula was unable to generate enough numbers and the density became flat and the gap between the two densities kept increasing. The total number of solutions found by Ramanujan's formula in the first 10,000 values of *a* were 1195. This was very small compared the numbers found by brute force: 65,085. Therefore, we concluded that while Ramanujan's formula does generate numbers satisfying the equation, it fails to generate enough density or count compared to the method of brute force.

DISCUSSION

In this paper we found solutions of following three Diophantine equations of the third power by brute force for the first ten thousand numbers of a.

$a^3 + b^3 + c^3 = d^2$ (quadratic equation or square equation)
$a^3 + b^3 + c^3 = d^3$ (cubic equation)
$a^3 + b^3 + c^3 = d^4$ (quartic equation or 4th power equation)

Here *a* is the largest number on the left-hand side of all the three Diophantine equations. We compiled all the solutions for the three equations into graphs showing density of solutions per hundred numbers of a and compared them. We found that while the quadratic equation and cubic equation showed a similar trend in density as the numbers became large (they rose rapidly at first and later the rise slowed down), the quartic equation's density was constant and constrained. The quartic equation's density at any point was much less than the density of the quadratic and cubic equations at that point.

We also compared the total number of solutions per combinations of *a*, *b* and *c* as *a* increased from 1 to 10,000. This graph fell rapidly at first in all the three cases and then stabilized. The fall was much larger for the fourth power equation compared to the cubic and quadratic equation.

We also found that the quadratic equation produced the greatest number of solutions followed by the cubic equation, whereas the quartic equation produced less than thousand solutions in our data.

We found numbers which we called perfect power taxicab numbers that are perfect squares, cubes or 4th powers, according to the equation, that could be expressed as a sum of three cubes in different ways. We compiled the perfect power taxicab numbers into histograms to find the count of numbers corresponding to various frequencies. We found that the number of perfect square taxicab numbers was much more than the number of perfect cube taxicab numbers. The number of perfect 4th power taxicab numbers was the least. Surprisingly, one perfect cube taxicab number

(970,299,000,000) could be expressed as the sum of three cubes in 42 different ways. In the cubic equation we were surprised to find that the count of numbers with frequency of 3 was more than the count of numbers with frequency of 2.

Finally, we generated solutions of the cubic equation given by Ramanujan's formula and compared its density with ours. We found that Ramanujan's formula could not match the density or count of solutions found by us using brute force.

When we compared our results with our hypothesis, we found that our hypothesis was correct in that the number of solutions rose with an increase in a for the square and the cubic equation. For the 4th power equation, however, the density fluctuated rapidly but stayed stable as a increased. We did not expect to see any taxicab number that could have a frequency of more than 20. However, in the cubic case we found a number (970,299,000,000) with the frequency of 42. Ramanujan's formula to find solutions of the cubic equation could only find a small fraction of the solutions that we found by brute force, and as the numbers grew larger, the gap became bigger. This trend was not what we expected in our initial hypothesis.

In this paper, we showed the trends of three Diophantine equations of the third power of the three input variables. We have merely explored a small area of number theory of Diophantine equations. An idea for a future research topic would be to analyze and compare the density of $a^3 + b^3 + c^3 + d^3 = e^3$ with the equations that we have studied. However, as the number of variables in the Diophantine equation increases, it takes up a lot of computation time to find all the solutions. We tried our Python programs with the four variable Diophantine equation. We ran the program for this equation for three days but the program did not even reach $a = 1000$. The reason for this is that the number of loops increases with the addition of another variable. So, if one Diophantine equation takes 1 hour to finish finding the solutions up to $a = 10{,}000$, the time taken to finish the same Diophantine equation with an added variable will increase by a factor of ten thousand (i.e. take 10,000 hours), and so on. Therefore, this is a good idea for a future research paper in which this equation can be tackled by better algorithms or by using powerful processors and distributed programming. We could also find solutions for all the above Diophantine equations till a equals one million to see how the density graphs change after $a = 10{,}000$. However, it would require faster computers and more efficient algorithms.

MATERIALS AND METHODS

In our paper we utilized the power of computer programming to create and evaluate our data. We wrote Python programs (Python version 3.7) for finding all our numbers for each Diophantine equation using brute force. By brute force we mean that we tried every combination of numbers to find the solutions to the equations. We redirected the output into a text file arranged as columns of *a*, *b*, *c* and *d*. We designed a separate Python program to parse that data into density per hundred numbers of *a*. To find the density of numbers we arranged all solutions in increasing order of *a*. Then another program found the difference of the line numbers of every consecutive multiple of 100 *a*s. In order to find curves that fit our data for easy comparison we used polyfit and poly1d functions that are available in the numpy scientific package of Python.

To find perfect power taxicab numbers, we created Python programs that would take every d and tried to find if the same d could be expressed in different ways. It did this by parsing our file and matching the d values.

To generate the solutions obtained by Ramanujan's equation we needed to find x and y such that the equation $(3x^2 + 5xy - 5y^2)^3 + (4x^2 - 4xy + 6y^2)^3 + (5x^2 - 5xy - 3y^2)^3 = (6x^2 - 4xy + 4y^2)^3$ could yield solutions within the range that we had selected: a ranging from 0 to 10,000 and $a \geq b$ and c. To accomplish this, we considered the three contributing terms: $3x^2 + 5xy - 5y^2$, $4x^2 - 4xy + 6y^2$ and $5x^2 - 5xy - 3y^2$. Each of these had to be positive integers and less than 10,000. So, we plotted the following inequations:

$$3x^2 + 5xy - 5y^2 > 0$$
$$3x^2 + 5xy - 5y^2 < 10000$$
$$4x^2 - 4xy + 6y^2 > 0$$
$$4x^2 - 4xy + 6y^2 < 10000$$
$$5x^2 - 5xy - 3y^2 > 0$$
$$5x^2 - 5xy - 3y^2 < 10000$$

Out of the above only the curves corresponding to $4x^2 - 4xy + 6y^2$ were bound and closed. From this we could determine the values of x needed to calculate the solutions to the cubic equation in the range of values that we had selected. We found that the x should range from -55 to +55. The plot of the other equations showed us that y should always be less than x. We implemented a Python program which we looped x and y from -55 to 55 and selected only those solutions for which a, b, c and d were positive and the largest of the three inputs is less than or equal to 10,000; the rest of the numbers were discarded.

We arranged all our data output by Python programs separated by '|', so it was easier to input this data into Microsoft Excel (version 1902) to create graphs which we could then analyze. Then, we arranged this data per hundred numbers of *a* and compared its density to that achieved by our data. We also used commands like 'grep' for finding a number in our data, 'sort' for sorting data in an order, 'wc' for counting lines, 'sed' for deleting or adding characters at every line. These commands were available to us as we operated on a Linux machine.

Since our programs were complex and computationally intensive, we purchased an instance on Amazon Web Services for computational time. We needed this to speed up the execution of our programs. While searching for numbers, we were careful about any defects in our data. Initially, there were several problems in our data which we fixed through

rigorous searching and debugging. We always kept a backup of old data by making new files at every stage.

Received: February 26, 2020
Accepted: September 19, 2020
Published: October 5, 2020

REFERENCES

1. Weisstein, Eric W. “Diophantine Equations.” From *MathWorld- A Wolfram Web Resource*, mathworld.wolfram.com/DiophantineEquation.html
2. Weisstein, Eric W. "Pythagorean Triple." From *MathWorld-A Wolfram Web Resource*, mathworld.wolfram.com/PythagoreanTriple.html
3. Bell, Eric Temple. *Men of Mathematics*. Simon & Schuster, 1986.
4. “Fermat’s Last Theorem” *Encyclopædia Britannica,* Encyclopædia Britannica, Inc., 13 Jun. 2020, https://www.britannica.com/science/Fermats-last-theorem
5. “Srinivasa Ramanujan.” *Encyclopædia Britannica,* Encyclopædia Britannica, Inc., 29 Jan. 2020, www.britannica.com/biography/Srinivasa-Ramanujan
6. Kanigel, Robert. *The Man Who Knew Infinity.* Washington Square Press, 1992.
7. Piezas, Tito III and Eric W. Weisstein. "Diophantine Equation--3rd Powers." From *MathWorld--A Wolfram Web Resource*, mathworld.wolfram.com/DiophantineEquation3rdPowers.html
8. Berndt, B. C. *Ramanujan's Notebooks, Part IV*. New York: Springer-Verlag, 1994
9. Dickson, L. E. History of the Theory of Numbers, *Vol. 2: Diophantine Analysis*. New York: Dover, 2005.
10. Burton, David. *Elementary Number Theory*. McGraw-Hill, 2017.
11. Weisstein, Eric W. "Taxicab Number." From *MathWorld--A Wolfram Web Resource*, mathworld.wolfram.com/TaxicabNumber.html

JOURNAL OF
EMERGING INVESTIGATORS

Article

Ladder Fuel Treatments Effect Burn Area of Forest Fires in Semi-Arid High Elevation Climates

Brian Schwartz[1], Keri Polevchak[1], Noah Eltzroth[1]
[1] duPont Manual High School, Louisville, Kentucky

SUMMARY

Forest fires destroy millions of acres, threaten populations, and are becoming increasingly severe. Efforts to reduce the starting of wildfires has led to increasing quantities of ladder fuels. Ladder fuels are any fuels that create a vertical continuity in a forest. When a fire does occur, these ladder fuels promote the spread of fires to the forest crown. We hypothesized that if the density of ladder fuel is reduced, the burn area in a semi-arid high-altitude forest will decrease. To test the hypothesis, a series of computer simulations were run testing 50 random ignition sites in a representative semi-arid high elevation forest with different levels of ladder fuel treatment using the FlameMap wildfire simulator with US Forest Service landscape and weather data. There was a large variation in the extent of the fires based on the ignition location, ranging from 13 to 22,506 acres. The study showed a statistically significant reduction in average total burn area after 10 days when 5% or more of the total acreage had been treated ($p < 0.01$). There were further reductions in burn area with increasing ladder fuel treatment from a 9% reduction with 5% treatment to a 51% reduction with 20% treatment ($p < 0.001$). This study supports the effectiveness of ladder fuel treatments in reducing burn area.

INTRODUCTION

In the United States, there are about 70,000 wildfires that affect millions of acres every year (1). The cost of wildfire suppression is almost one billion dollars a year (1). Over time, longer forest fires seasons and drought around the world have stretched limited wildfire fighting resources (1). The extensive property damage from forest fires, increasing cost of fighting wildfires, and limited firefighting resources make techniques which reduce wildfire damage even more valuable.

Forest fires are most common in the southwestern United States due to its semi-arid high elevation climate and large ponderosa pine forests. These forests are very homogenous, allowing forest fires to easily spread over a vast area (2). Before the 1820s, fires were smaller and burned less area than today (1). These fires occurring frequently in the semi-arid ponderosa forests (2). Through human intervention, smaller fires became less frequent as they were put out to protect the people who had settled in the region (3). Since there were fewer fires burning, fallen tree branches and other underbrush began to build up over the next several decades (3). One reason the underbrush build-up increased even more rapidly in the 1950s was because the United States Forest Service created Smokey the Bear to urge citizens to stop and prevent fires (4). Efforts to extinguish fires led to an excess of fuels in the forest (3). Due to the abundance of built-up fuels, when a forest fire did occur, the fires became large and uncontrollable (3). The massive increase in large fires also led to an outbreak of scientific research in the field of science ecology (3).

It is currently hypothesized inside the scientific community that these fuels cause more fires to become crown fires and therefore burn a larger area (4). Forest fires spread across horizontal and vertical continuities (5). Horizontal continuity is an unbroken existence of fuels in the horizontal direction. Ladder fuels create a vertical continuity of fuels in the vertical direction (6) and allow fires to spread via the tops of trees becoming crown fires which are the most dangerous and fastest moving type of fire (5). This fact is of great significance as crown fires are hazardous, difficult to extinguish, and most importantly, the number of crown fires is increasing (3). If ladder fuels are found to be the cause of ground fires becoming crown fires, decreasing ladder fuels can be used to prevent the increasing amount of crown fires.

Fuel treatments are any method that remove flammable materials in either the vertical or horizontal continuity from the forest environment to reduce the number and extent of forest fires (3). The most frequently used fire treatments are prescribed burns and thinning to prevent fires from spreading up trees and becoming crown fires (3). This project will provide information on how fuels can be treated to better protect people, wildlife, and property from forest fires.

Cochrane et al. found that there is a direct correlation between the burn area of forest fires and the density of both horizontal and vertical continuities (7). With this discovery, we hypothesized that if the density of ladder fuels decreases through clearing, the burn area will be reduced in the tested environments. This decrease is due to the belief that vertical continuity will influence the burn area of forest fires. The null hypothesis is true if the reduction of ladder fuels does not reduce the burn areas in the tested environments. It can be inferred in that situation that horizontal continuity affects the burn area of forest fires, but vertical continuity does not affect the burn area.

The study was conducted by running a simulator in

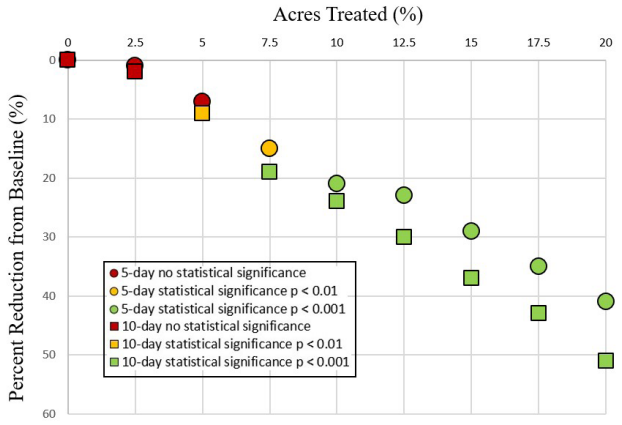

Figure 1: Five -day and ten-day measurements of average acres burned at different percentage of land pretreated levels (0-20% treated). Greater slope in 10-day vs. 5-day data shows greater impact of more extensive ladder fuel reduction with time. The difference between the two data sets demonstrates the effect of time on fire spread.

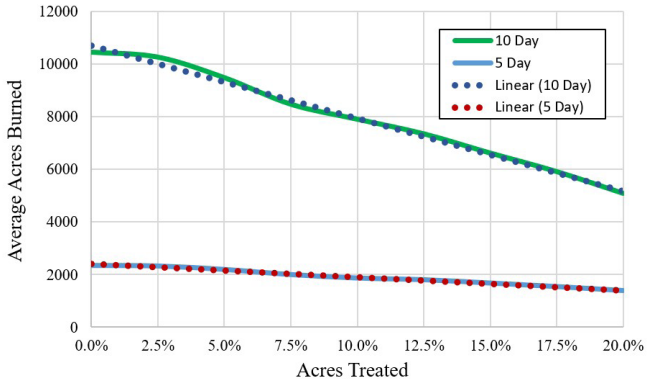

Figure 2: Ladder fuel treatment results in a significant reduction in acres burned over five- and ten-days. Statistical significance (p < 0.01) is shown beginning at 7.5% treatments with 5-day data and 5% with 10-day data. p < 0.01 is represented by yellow markers and p < 0.001 is represented by green markers. (For each treatment group, n = 50.)

different semi-arid high-elevation environments with different levels of ladder fuels. The burn area will be found for each of the simulated fires. This data was evaluated to find if there is a statistical difference between burn area when there is a difference in ladder fuels.

RESULTS

Using FlameMap, ladder fuels were reduced on varying amounts of the environment. Nine different percentages, from 0% to 20% of the total acreage, was treated to remove ladder fuels. The land treated was determined by starting at the lowest elevation. 50 simulated fire ignitions at each of the 9 different levels of fire ladder treatment produced 450 data points for 5- and 10-day data. There was a wide variation in total burn areas in the 10-day data from as little as 13 acres to as much as 22,506 acres depending on the location of ignition. The range was smaller for the 5-day data between only 2 and 5,800 acres. The duration of the fires also impacted the extent of burn area. The untreated baseline 10-day average burn area was more than 4 times the burn area of the 5-day measurement. The effect of fuel treatment was magnified with the longer duration simulations. Untreated 5-day burn area was 3.7 times greater than the heavily treated 20% landscape, but after 10 days was 4.4 times greater (**Figure 1**). The extent of area burned, and effectiveness of ladder fuel treatment varied greatly depending on the ignition location and the duration of the fire.

This study evaluated the effect of ladder fuel treatment on reducing burn area in semi-arid high elevation forest. The average burn area progressively decreased from baseline to 20% treatment in both 5- and 10-day data (**Figure 1**), showing a clear reduction in burn area with increases in ladder fuel treatments. One-way ANOVA analysis was done for both 5- and 10-day data. The 10-day results were statistically significant (p-value 1.1 x 10-5). ANOVA results from 5-day data had a p-value of 0.03 which did not meet the alpha of 0.01 used for this study. With the 5-day data, a statistically significant difference was seen beginning with 7.5% treated and above when compared with the untreated forest. There was also a statistically significant drop (p < 0.01) in burn area between each incremental increase in ladder fuel treatment beyond the first 2.5% in the 10-day simulations. In the 5-day data, statistical significance (p < 0.01) was seen between 7.5% to 10%, 12.5% and 15%, and 17.5% and 20%.

There was an inverse relationship between the average acreage burned and percent of landscape treated (**Figures 1 & 2**). A comparison of 10-day and 5-day data showed a greater reduction in burn area in total acres burned and a greater reduction in the percentage of acres burned with the longer simulation time (**Figure 1**). For example, with the 20% treatment group, average burn area for 10 days was 5,365

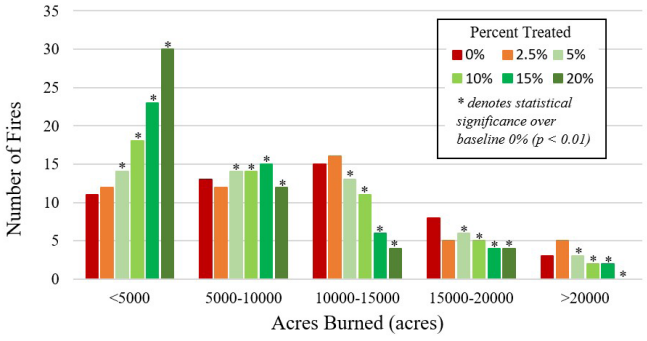

Figure 3: Percent of land pretreated organized by number of acres burned. ANOVA analysis of 10-day data showed a statistically significant variation in the different treatment levels. Green shaded data is significantly less (p < 0.01) then baseline (0%) group. The graph demonstrates the extent that increasing ladder fuel treatments constrains wildfires to smaller areas. The highest levels as ladder fuel treatments increase, there is a progressive reduction in forest fire size. Only the most extensive fuel treatments completely prevent the largest burn areas (zero fires of greater than 20,000 acres with 20% treatment).

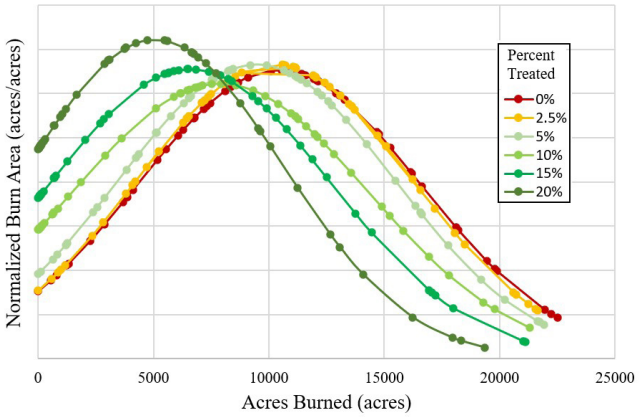

Figure 4: Ten-day data normalized to a Gaussian distribution showing acres burned for each percent of land treated. There is a shift to the left with higher levels of pretreatment showing reduced average burn areas. Narrowing of the distribution curve with increasing fuel treatment shows a reduction in the variability of wildfires with greater fuel reduction.

acres less and 51% less than baseline, while the 5-day data was 971 acres less and 41% less than baseline (**Figure 1**).

The fire data was divided into groups of 5,000 acres burned (**Figure 3**). With higher levels of ladder fuel treatment, the fires were less extensive. At 20% treatment, 60% of fires were less than 5,000 acres after 10 days. With little treatment of 0% or 2.5%, 20 percent or more of fires were greater than 15,000 acres.

The 10-day data was normalized to a Gaussian distribution (**Figure 4**). The graph shows a shift to the left as the mean burn area is reduced 50% with highest levels of pretreatment. Only at the higher levels of ladder fuel treatment tested, particularly at 20%, is there an increase in the height of the of the curve and narrowing of curve width indicating the variation in burn areas is reduced. The reduction in standard deviation data with 20% ladder fuel treatment confirms the reduced variability.

DISCUSSION

In this study, through utilization of the FlameMap simulator, a statistically significant reduction in burn area in semi-arid high elevation forest fires with ladder fuel treatment in 5% or more of the acreage. The extent of reduction in burn area correlated with the degree of pretreatment to reduce ladder fuels. The simulated comparison between untreated and treated areas was made using identical wind, weather, and moisture conditions. There was also a marked variation in the size of forest fires depending on the location of ignition, duration of fire, and extent of ladder fuel treatment. These findings illustrate the effectiveness of ladder fuels reduction on forest fire containment.

Ladder fuel reduction treatment on 5% of the forest produced a statistically significant reduction in burn area of 9% after 10 days ($p < 0.01$) and treating 20% of the land reduced the burn area 51% ($p < 0.001$). The study refuted the null hypothesis, demonstrating a significant reduction in burn area with increasing ladder fuel reduction.

These results are particularly important with the increasing severity and cost of forest fires. While increasing costly firefighting resources are needed, efforts to limit the severity and destructiveness of forest fires before burning begins deserves additional consideration. This study shows that ladder fuel reduction can have a significant impact on the speed and size of fires. The cost of ladder fuel reduction varies greatly depending on access to mechanical equipment, topography, and remoteness (8). Estimates for the western US range from $1,000 to more than $3,000 per acre (8). Even using the low-end estimate, the cost would be greater than $350 million dollars just to treat 5% of the test landscape which is an exceedingly small portion of the US semi-arid high elevation forest. While treatment is not economically viable on a national level, ladder fuel reduction around the intersection of forest with urban areas could be useful and economical. Judicious use of ladder fuel treatments could reduce some of the annual one billion dollars spent on fighting forest fires (1).

Our study was limited by several factors including the testing of ladder fuel treatment to reduce vertical spread without the combination of horizontal barriers. This limitation, as well as additional questions, can be addressed by further research. The simulation was specifically based on environmental data from the 2016 peak fire season (9). Further simulated studies using a wider variety of environmental conditions such as other seasons of the year and data from other years than 2016 could further support the effectiveness of ladder fuel reduction. This research looked at seven million acres of semi-arid high-altitude forest, but this is only a small portion of this type of forest in the United States and might not be representative of other areas. Running the simulation software over a wider range of forests including other types of forests would determine the generality of this study's results. This study-based treatment on altitude but it may be more realistic to base areas of treatment around access roads which would be critical for large scale ladder fuel reduction. One significant advantage of using a simulated comparison is the ability to adjust and control the environmental conditions. Further studies with simulated environmental conditions anticipated by climate change could provide insight on the future utility of ladder fuel treatment.

This study demonstrated the effectiveness of ladder fuel treatment in semi-arid high elevation forest and this technique maybe beneficial when combined with horizontal barriers to reduce the destructiveness of forest fires at the intersection of forest with urban areas.

MATERIALS AND METHODS

The FlameMap software was downloaded from the US Forestry website. The FlameMap Simulator is known for its accuracy in predicting forest fires (5, 6, 10-13). This simulator is also currently in use by the American government and

forest firefighters (5). The FlameMap Simulator allows for a variety of input environments and an accurate simulator for determining the burn area and category for each test.

Environmental Data

The landscape data was retrieved from landfire.gov (9). A rectangular 7,083,100 acre landscape from latitude 37° longitude -106° to latitude 36° longitude -105° was used. Landscape files were the latest available from 2019. Historical weather data from forestry service weather stations was used from July and August of 2016.

The FlameMap simulation used generated gridded wind from the available weather input data. Fuel moisture conditioning was started 10 days prior to the ignition of the fire. Ember spot probability was set to 0.10 and crown fire calculation was Scott/Reinhardt. The ignition start points were determined randomly using the Microsoft Excel random number generator to determine a matrix location on the landscape.

Pretreatment Parameters

The landscape was tested using baseline state and various degree of pre-treatment. The baseline was determined by running the FlameMap simulation utilizing weather, topography, and fuel data from US Forestry website for July and August 2016. The ladder fuels were reduced at the lowest levels of elevation first as these would be the least costly to effect. Pre-treatments at 2.5%, 5%, 7.5%, 10%, 12.5%, 15%, 17.5%, and 20% were used. Previous studies in other environments have shown using ladder fuel reduction treatment in 8% of the area creates a significant reduction in burn area (14).

Statistical Analysis

Fifty ignition points were tested at baseline and treated landscapes. The data for burn area was collected at both 5- and 10 -days. The data was then collected and analyzed using Microsoft Excel. Both one-way ANOVA and two-tailed t-test were run, in order to determine if there was statistical significance between the different environments. In addition, according to Devore, the t-test was created to determine the existence of statistical significance and that the bright-line for statistical significance for computer studies is 0.01 (15).

ACKNOWLEDGEMENTS

We thank Ms. Alesia Williams (duPont Manual High School) for her advice on writing a research paper. Editing assistance and support from Mrs. Ann Schwartz and Dr Robert Schwartz is also appreciated.

Received: June 18, 2020
Accepted: September 21, 2020
Published: October 5, 2020

REFERENCES

1. *National Interagency Fire Center*. National Interagency Coordination Center. 2017.
2. Stafford, Hugh D., and Jens T. Stevens. *Natural Range of Variation (NRV) for Yellow Pine and Mixed Conifer Forests in the Bioregional Assessment Area, Including the Sierra Nevada, Southern Cascades, and Modoc and Inyo National Forests*. Vallejo, CA, U.S. Department of Agriculture, Forest Service, Pacific Southwest Research Station, 2013.
3. Omi, Philip N. *Forest Fires: A Reference Handbook*. Santa Barbara, CA, ABC-CLIO, 2005.
4. Stine, Peter, et al. The Ecology and Management of Moist Mixed-Conifer Forests in Eastern Oregon and Washington: a Synthesis of the Relevant Biophysical Science and Implications for Future Land Management. Portland, OR, *U.S. Department of Agriculture, Forest Service, Pacific Northwest Research Station*, 2014, doi:10.2737/pnw-gtr-897.
5. Finney, Mark A. "FARSITE: Fire Area Simulator-Model Development and Evaluation." *U.S. Department of Agriculture, Forest Service, Rocky Mountain Research Station*. 2004, Ogden, UT, doi:10.2737/rmrs-rp-4.
6. Reinhardt, Elizabeth D., and Matthew B. Dickinson. "First-Order Fire Effects Models for Land Management: Overview and Issues." *Fire Ecology*, vol. 6, no. 1, 2010, pp. 131–150, doi:10.4996/fireecology.0601131.
7. Cochrane, M. A., et al. "Estimation of Wildfire Size and Risk Changes Due to Fuels Treatments." *International Journal of Wildland Fire*, vol. 21, no. 4, 2012, pp. 357-367, doi:10.1071/wf11079.
8. *Fuel Reduction and Forest Restoration Plan for the Lake Tahoe Basin Wildland Urban Interface*. Steve Holl Consulting. 2007. Section 7: Cost Estimates, 7.1-7.6.
9. *Landfire*. U.S. Department of the Interior, U.S. Department of Agriculture Forest Service.
10. Rothermel, Richard C. "A Mathematical Model for Predicting Fire Spread in Wildland Fuels." *U.S. Department of Agriculture, Forest Service, Rocky Mountain Research Station*. 1972, Ogden, UT.
11. Rothermel, Richard C. "Predicting Behavior and Size of Crown Fires in the Northern Rocky Mountains." *U.S. Department of Agriculture, Forest Service, Rocky Mountain Research Station*. 1991, Ogden, UT. doi:10.2737/int-rp-438.
12. Van Wagner, C. E. "Conditions for the Start and Spread of Crown Fire." *Canadian Journal of Forest Research*, vol. 7, no. 1, 1977, pp. 23-34. doi:10.1139/X77-004.
13. Van Wagner, C. E. "Prediction of Crown Fire Behavior in Two Stands of Jack Pine." *Canadian Journal of Forest Research*, vol. 23, no. 3, 1993, pp. 442–449, doi:10.1139/X93-062.
14. Syphard, Alexandra D., et al. "Simulating Landscape-Scale Effects of Fuels Treatments in the Sierra Nevada, California, USA." *International Journal of Wildland Fire*, vol. 20, no. 3, 2011, pp. 364-383. doi.:10.1071/ wf09125.

15. Devore, Jay L. *Probability and Statistics for Engineering and the Sciences*. Toronto, Ontario, Brooks/Cole Publishing Company, 1987, pp. 285-286.

JOURNAL OF
EMERGING INVESTIGATORS

Article

The Effects of Barley Straw (*Hordeum vulgare*) Extract and Barley Straw Pellets on Algal Growth and Water Quality

Elaina McHargue[1], Chelle Gillan[1]
[1]Central City High School, Central City, NE

SUMMARY

In recent years, harmful algal blooms have increased in both frequency and intensity worldwide. This is a growing concern because many algae species can clog agricultural irrigation systems, make potable water unfit for consumption, and release toxins that can be dangerous to human and animal health. The purpose of this study was to test a natural solution for inhibiting algal growth that does not expose animals to potentially harmful chemicals. This experiment tests the effects of barley straw extract (BSE) and barley straw pellets (BSP) on algal growth and water quality. We predicted that both treatments would have significant effects on algal growth and water quality. We added BSE and BSP to containers of water placed them in a temporary greenhouse. Then, we conducted several different types of tests to determine the amount of algal growth as well as the quality of the water over a course of 25 days. The results showed some significant differences between the treatments on certain testing days with the transmittance, dissolved oxygen (DO), and carbonate tests; however, the results were not conclusive enough to reject the null hypothesis that neither treatment would have significant effects on algal growth and water quality. Research of this type has value because it is important to protect human and animal health by providing clean water sources.

INTRODUCTION

Livestock, outdoor pets, and wildlife all rely on outdoor water sources that are susceptible to algal blooms during the warm summer months. In recent years, harmful algal blooms (HABs) have increased in both frequency and intensity worldwide (1). There are many reasons to be concerned about the growing trend of HABs in our society today. First, many species of algae can create unappealing odors and tastes which minimize the water intake of livestock and other animals in addition to making it unfit for human consumption. Second, as another concern with agriculture, filamentous algae can clog pumps, screens, and emitters in agricultural irrigation systems (2). Third, the decaying process of HABs can cause eutrophication, or the depletion of oxygen in a lake or pond. This can lead to the death of many organisms, such as fish (3). Finally, many strains of algae and cyanobacteria, also called blue-green algae, can produce harmful toxins. There are two general types of toxins that are produced by algae and cyanobacteria: neurotoxins and hepatotoxins. Neurotoxins are rapid acting, deadly toxins that influence nerve cells, and their effects on behavior and movement can usually be observed within minutes. Hepatotoxins, on the other hand, are slower acting toxins that damage the liver, causing vomiting, changes in heart rate, and even death (4). These effects are usually observed within a few hours (4).

Not only do HABs cause issues directly by the reasons listed above, they can also cause issues indirectly through environmentally harmful algicide treatments (2). Such treatments can affect animal health, and, since most are not species specific in their inhibitory effects, can eliminate higher plants in addition to algae. For these reasons, many people have attempted to mechanically control algae by way of raking, cutting, or harvesting; however, this method is costly and ineffective because fragments of algae still remain and regrow rapidly (5). The goal of this study was to test a plant-based solution for inhibiting algal growth that will not introduce potentially harmful chemicals into the water.

We tested barley straw (*Hordeum vulgare*) in this study. Barley straw is an algistatic solution that is used to inhibit algal growth in lakes and ponds (6). As the barley material decomposes, it releases free radicals, such as hydrogen peroxide, that carry out oxidative damage to algal cells (7). The initial studies on barley straw as an algal growth inhibitor were conducted in England in the 1990s, and it has grown in popularity across the United States in recent years (6). Most commonly, barley straw is applied to lakes and ponds by placing bales of it in nets and floating them in the water. For the purposes of this experiment, barley straw was used in both a liquid extract and a solid pellet form. This enabled us to test smaller amounts of water in a more controlled environment.

To summarize, an increasing number of HAB occurrences have been observed globally in recent years. Such occurrences can have detrimental effects on human and animal health, agriculture, and recreation in lakes and ponds. Many methods of inhibiting algal growth have been tested with limited amounts of success. The purpose of this experiment was to discover the effects of BSE and BSP on algal growth and water quality. The experimental design was made with the intention of testing the effectiveness of barley straw in an authentic setting during warm summer months in which the conditions are ideal for airborne algal spores to enter water sources and rapidly reproduce. Most freshwater green algae species can spread in this way (8). We hypothesized that both the BSE and the BSP would have significant effects on algal growth and water quality; however,

data from experimentation in the study did not indicate that either treatment had significant effects on algal growth or water quality. These results indicate that barley straw may not be a cost-efficient solution to reducing algal growth, as it is not effective in all cases.

RESULTS

To determine the effects of BSE and BSP on algal growth and water quality, we placed six replicates of each experimental group in five-liter buckets and randomly assigned them spots in a temporary greenhouse. The experimental groups consisted of BSE and well water, BSP and well water, and plain well water as the control treatment. Over a course of 25 days, we conducted spectrophotometer, hemocytometer, and water quality tests on each of the samples by monitoring nine different indicators of algal growth and water quality.

We measured the percentage of light transmittance because it measures water clarity, and, as algae grows, water becomes murkier, resulting in lower transmittance levels. The light transmittance in the control (C) samples went from an average of 100% transmittance on the first day to an average of 72.3±1.2% transmittance on the twenty-fifth day. The BSE samples went from an average of 100% light transmittance on the first day to an average of 78.2±3.6% transmittance on the twenty-fifth day. The barley straw pellet samples went from an average of 100% light transmittance on the first day to an average of 68.1±2.9% transmittance on the twenty-fifth day. A single-factor ANOVA test followed by a Tukey-Kramer test showed a significant difference when comparing the following: control to BSP and BSE to BSP on day 6 (P<0.001), and control to BSP and BSE to BSP on day 11 (P<0.05). Therefore, BSP did appear to significantly affect algal growth on days 6 and 11, as measured by transmittance (**Figure 1**).

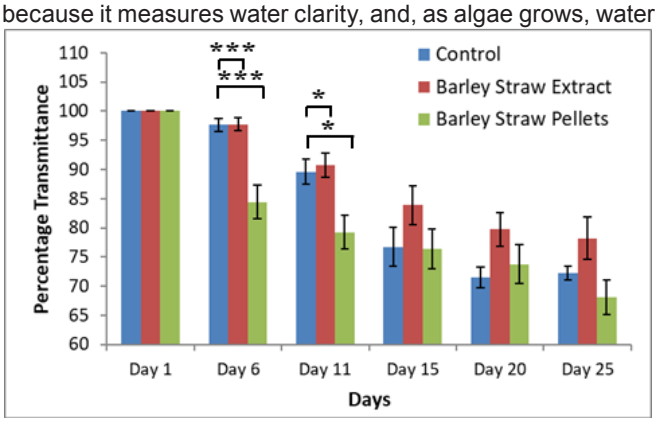

Figure 1. Percent transmittance decreased over time for all treatments. We placed samples (n=6) in a greenhouse for 25 days during the month of August. Data points indicate the mean transmittance (ppm). Asterisks indicate significant difference (*P<0.05, ***P<0.001).

Because it is essential to the survival of fish and other aquatic organisms, we measured DO in the samples (9). All of the samples increased in DO levels until day 15, when they started to trend downward. The control samples went from an average of 8.10±0.10 mg/L on the first day to an average of 9.65±0.24 mg/L on the twenty-fifth day. The BSE samples went from an average of 8.17±0.11 mg/L on the first day to an average of 10.17±0.23 mg/L on the twenty-fifth day. The barley straw pellet samples went from an average of 8.17±0.12 mg/L on the first day to an average of 10.68±0.22 mg/L on the twenty-fifth day. A single-factor ANOVA test followed by a Tukey-Kramer test showed a significant difference when comparing the following: control to BSP on days 15 and 25 (P<0.05) (**Figure 2a**). All of the treatments showed a negative correlation to some degree between DO and transmittance levels (C: -0.73 [strong], BSE: -0.67 [moderate], BSP: -0.70 [strong]), indicating that DO increased as transmittance decreased (**Figure 2b-d**). Therefore, BSP did appear to significantly affect water quality on days 15 and 25, as

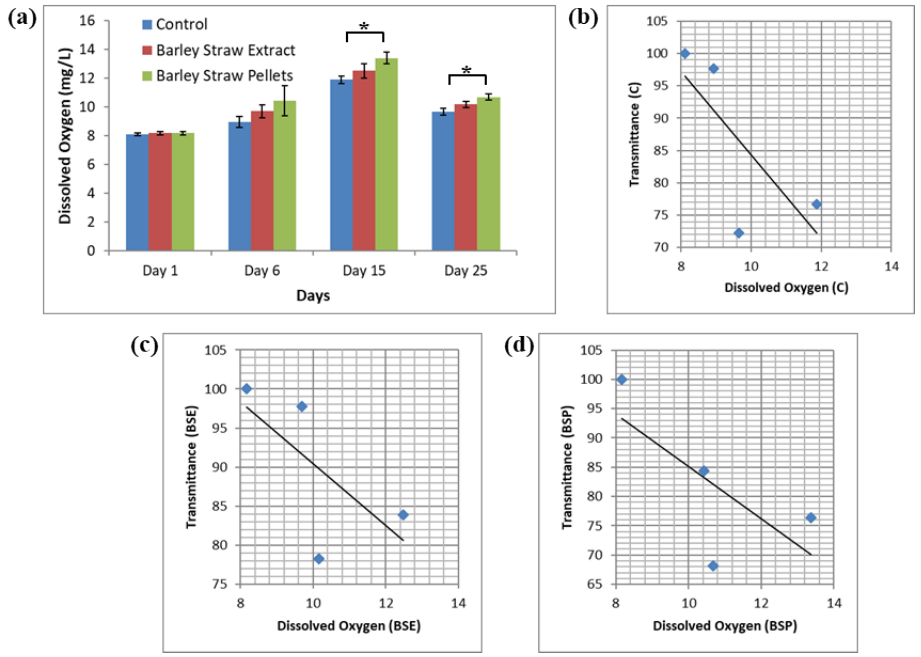

Figure 2. DO trended upward through day 15 and then decreased for all treatments, and all treatments' mean DO showed a negative correlation to transmittance to some degree. (**a**) We placed samples (n=6) in a greenhouse for 25 days during the month of August. Data points indicate the mean DO (mg/L), and error bars denote standard error. Asterisks indicate significant difference (*P<0.05). (**b-d**) Data points indicate mean DO compared to mean transmittance. Best fit line denotes correlation.

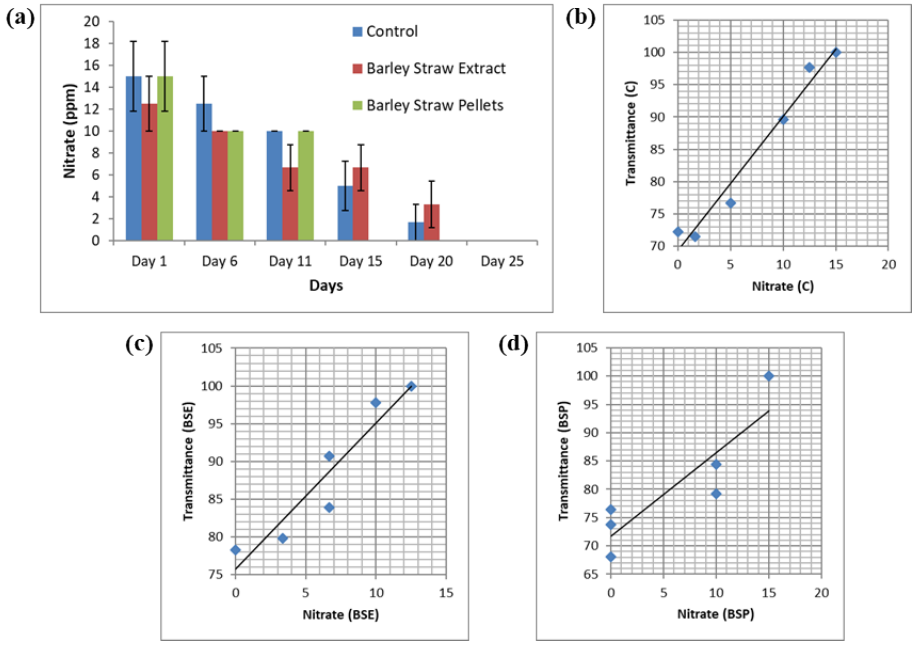

Figure 3. Nitrate levels decreased over time for all treatments and all treatments' mean nitrate levels showed a positive correlation to transmittance to some degree. (**a**) We placed samples (*n*=6) in a greenhouse for 25 days during the month of August. Data points indicate the mean nitrate levels (ppm), and error bars denote standard error. (**b-d**) Data points indicate mean nitrate levels compared to mean transmittance. Best fit line denotes correlation.

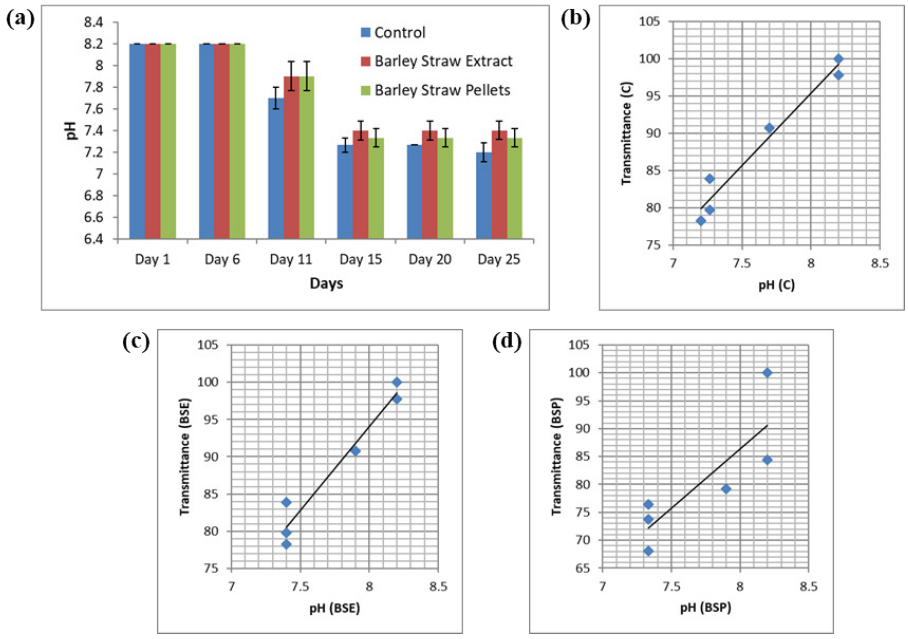

Figure 4. (**a**) **pH levels decreased over time for all treatments, and all treatments' mean pH levels showed a positive correlation to transmittance to some degree.** (**a**) We placed samples (*n*=6) in a greenhouse for 25 days during the month of August. Data points indicate the mean pH levels. (**b-d**) Data points indicate mean pH levels compared to mean transmittance. Best fit line denotes correlation.

measured by DO.

We measured nitrate levels because algae uses nitrates as nutrients to grow (1). The nitrate levels in the control samples went from an average of 15±3.2 ppm on the first day to an average of 0±0 ppm on the twenty-fifth day. The BSE samples went from an average of 13±2.5 ppm on the first day to an average of 0±0 ppm on the twenty-fifth day. The barley straw pellet samples went from an average of 15±3.2 ppm on the first day to an average of 0±0 ppm on the twenty-fifth day. A single-factor ANOVA test showed no significant difference when comparing the nitrate levels of each of the treatments (**Figure 3a**). All of the treatments showed a positive correlation to some degree between nitrate and transmittance levels (C: 0.98 [very strong], BSE: 0.94 [very strong], BSP: 0.88 [strong]), indicating that as nitrate levels increased, transmittance also increased (**Figure 3b-d**). Therefore, BSE and BSP did not appear to significantly affect water quality, as measured by nitrates.

Because they can indicate if water is changing chemically, we measured pH levels in the samples. The pH levels in the control samples went from an average of 8.2±0 on the first day to an average of 7.2±0.1 on the twenty-fifth day. The BSE samples went from an average of 8.2±0 on the first day to an average of 7.4±0.1 on the twenty-fifth day. The barley straw pellet samples went from an average of 8.2±0 on the first day to an average of 7.3±0.1 on the twenty-fifth day. A single-

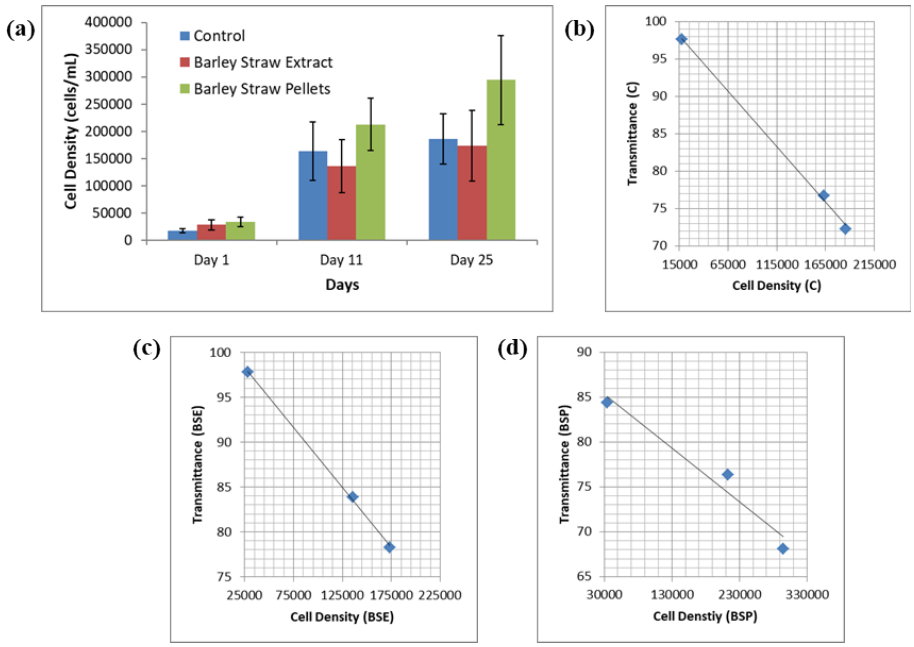

Figure 5. Cell density increased over time for all treatments, and all treatments' mean cell density showed a negative correlation to transmittance to some degree. (**a**) We placed samples (*n*=6) in a greenhouse for 25 days during the month of August. Data points indicate the mean cell density (cells/mL) (**b-d**) Data points indicate mean cell density compared to mean transmittance. Best fit line denotes correlation.

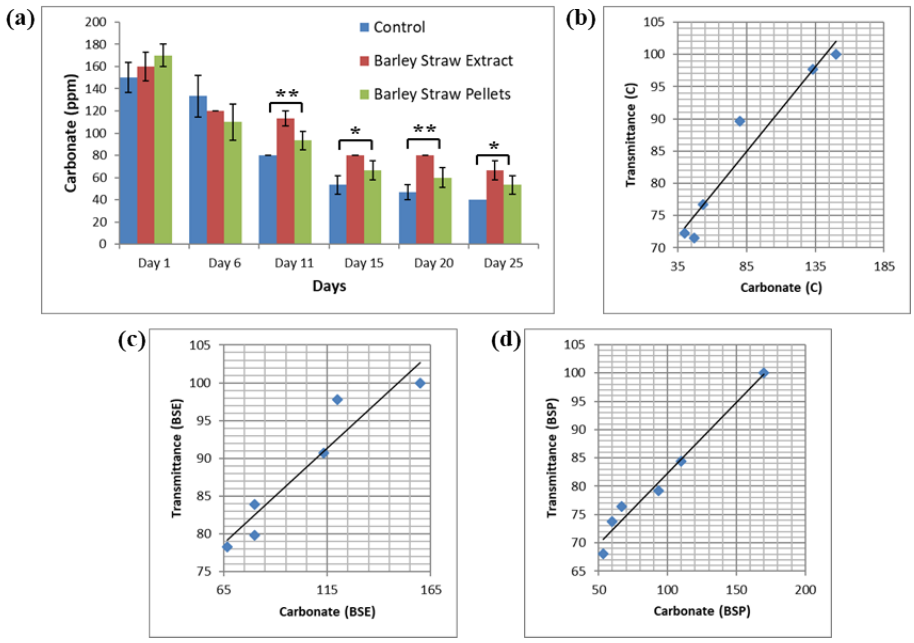

Figure 6. (**a**) Carbonate levels decreased over time for all treatments. (**b-d**) All treatments' mean carbonate levels showed a positive correlation to transmittance to some degree. (**a**) We placed samples (*n*=6) in a greenhouse for 25 days during the month of August. Data points indicate the mean carbonate levels (ppm). Asterisks indicate significant difference (**P*<0.05, ***P*<0.01). (**b-d**) Data points indicate mean carbonate levels compared to mean transmittance. Best fit line denotes correlation.

factor ANOVA test showed no significant difference when comparing the pH levels of each of the treatments (**Figure 4a**). All of the treatments showed a positive correlation to some degree between pH and transmittance levels (C: 0.98 [very strong], BSE: 0.98 [very strong], BSP: 0.83 [strong]), indicating that as pH levels increased, transmittance also increased (**Figure 5b-d**). Therefore, BSE and BSP did not appear to significantly affect water quality, as measured by pH.

We measured cell density because it quantifies the amount of algae that is growing in the water. The cell density levels for the control samples went from an average of 17,667±3,555.9 cells/mL on the fifth day to an average of 186,000±45,959.4 cells/mL on the twenty-fifth day. The BSE samples went from an average of 28,000±8,869.4 cells/mL on the sixth day to an average of 173,333±64,633.7 cells/mL on the twenty-fifth day. The barley straw pellet samples went from an average of 34,000±8,869.4 cells/mL on the fifth day to an average of 294,400±81,817.9 cells/mL on the twenty-fifth day. A Single-factor ANOVA test showed no significant difference when comparing the cell density levels of each of the treatments (**Figure 5a**). All of the treatments showed a negative correlation to some degree between cell density and transmittance (C: -0.99 [very strong], BSE: -0.99 [very strong], BSP: -0.98 [very strong]), indicating that as cell density increased, transmittance decreased (**Figure 5b-d**).

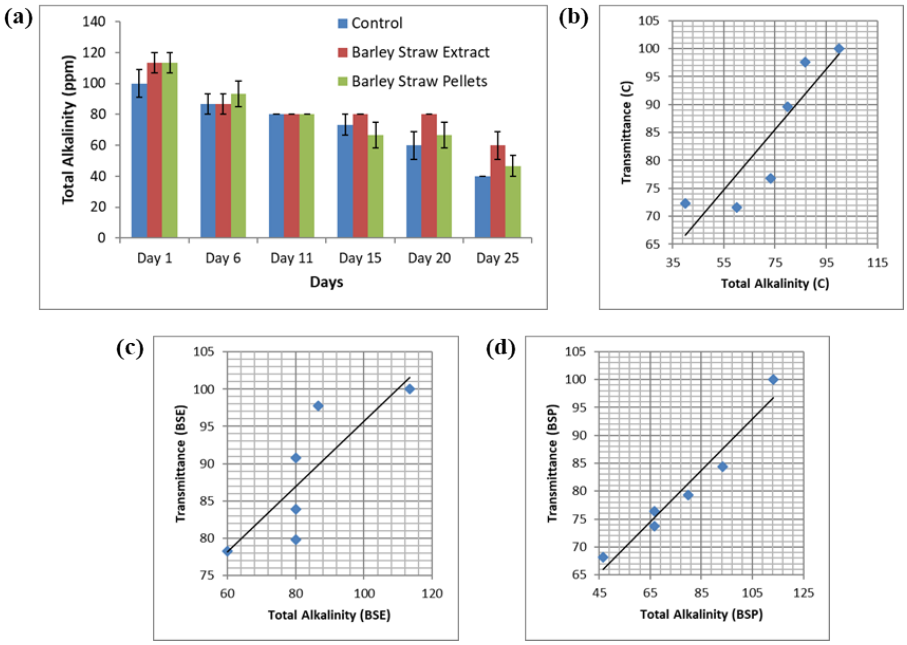

Figure 7. Total alkalinity levels decreased over time for all treatments, and all treatments' mean total alkalinity levels showed a positive correlation to transmittance to some degree. (**a**) We placed samples (*n*=6) in a greenhouse for 25 days during the month of August. Data points indicate the mean total alkalinity levels (ppm). (**b-d**) Data points indicate mean total alkalinity levels compared to mean transmittance. Best fit line denotes correlation.

Therefore, BSE and BSP did not appear to significantly affect algal growth, as measured by cell density.

Next, we measured carbonate levels because they can indicate water hardness (10). The carbonate levels for the control samples went from an average of 150±13.4 ppm on the first day to an average of 40±0 ppm on the twenty-fifth day. The BSE samples went from an average of 160±12.6 ppm on the first day to an average of 67±8.4 ppm on the twenty-fifth day. The barley straw pellet samples went from an average of 170±10.0 ppm on the first day to an average of 53±8.4 ppm on the twenty-fifth day. A Single-factor ANOVA test followed by a Tukey-Kramer test showed a significant difference when comparing the following: control to BSE on days 11-25 (P<0.01 on days 11 and 20, P<0.05 on days 15 and 20) (**Figure 6a**). All of the treatments showed a positive correlation to some degree between carbonate and transmittance levels (C: 0.96 [very strong], BSE: 0.95 [very strong], BSP: 0.99 [very strong]), indicating that as carbonate levels increased, transmittance also increased (**Figure 7b-d**). The control samples showed lower carbonate levels than the BSE and the barley straw pellet samples throughout the majority of the testing days, and they were significantly lower than the BSE samples on days 11-25.

Because it helps to balance pH levels, we also took total alkalinity into consideration. The total alkalinity levels of the control samples went from an average of 100±8.9 ppm on the first day to an average of 40±0 ppm on the twenty-fifth day. The BSE samples went from an average of 113±6.7 ppm on the first day to an average of 60±8.9 ppm on the twenty-fifth day. The barley straw pellet samples went from an average of 113±6.7 ppm on the first day to an average of 47±6.7 ppm on the twenty-fifth day. A Single-factor ANOVA test showed no significant difference when comparing the total alkalinity levels of each of the treatments (**Figure 7a**). All of the treatments showed a positive correlation to some degree between total alkalinity and transmittance levels (C: 0.89 [strong], BSE: 0.82 [strong], BSP: 0.9 [very strong]), indicating that as total alkalinity levels increased, transmittance also increased (**Figure 7b-d**). Therefore, BSE and BSP did not appear to significantly affect water quality, as measured by total alkalinity.

We monitored nitrite levels because high amounts of nitrite can disrupt oxygen transport (11). Only one measurable amount of nitrite was detected. This was with the barley straw pellet samples on day 1 (2±1.7 ppm). A Single-factor ANOVA test showed no significant difference when comparing the nitrite levels of each of the treatments. Neither the control nor the BSE showed any correlation between nitrite and transmittance levels, but the BSP showed a positive correlation of 0.87 (strong). Therefore, BSE and BSP did not appear to significantly affect water quality, as measured by nitrites.

Finally, we measured free chlorine levels because they can inactivate certain bacteria and viruses (12). Only one measurable amount of free chlorine was detected. This was with the control samples on the twentieth day (0.1±0.1 µg/L). A Single-factor ANOVA test showed no significant difference when comparing the free chlorine levels of each of the treatments. Neither the BSE nor the BSP showed any correlation between nitrite and transmittance levels, but the control showed a negative correlation of -0.50 (moderate). Therefore, BSE and BSP did not appear to significantly affect water quality, as measured by free chlorine.

DISCUSSION

This study was performed to understand how BSE and BSP affect algal growth and water quality. The results showed

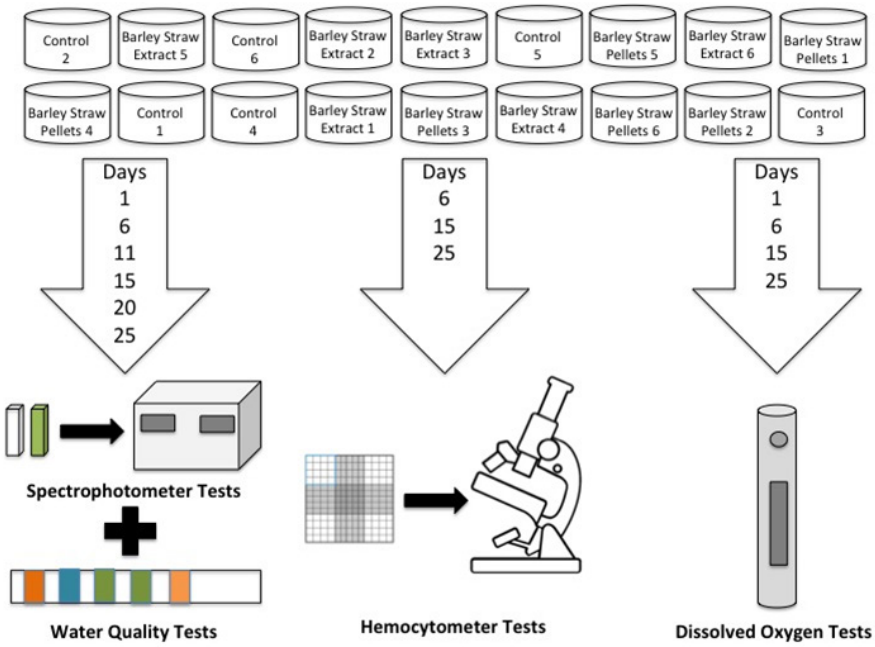

Figure 8. Procedure Diagram. Pictured is the layout of the samples in the greenhouse. The arrows going down show on which days each type of test was conducted.

statistically significant differences between the treatments on certain testing days with the transmittance, DO, and carbonate tests; however, the results were not conclusive enough across all tests to reject the null hypothesis that neither treatment will have significant effects on algal growth and water quality. There was, however, a correlation found when comparing the nitrate, DO, pH, cell density, carbonate, and total alkalinity tests to the transmittance tests. This suggests that the lack of significant differences between the treatments was not due to an error in experimentation, as the tests correlated with each other.

Each test showed its greatest change from days 6 to 15. This may indicate that days 6 to 15 were when the algal growth hit its peak. On day 15 for the BSP and 25 for the BSE and the control, the nitrate levels dropped down to zero, which likely means that the algae could no longer use that resource to grow. The BSP reached this point sooner than the other two treatments. It is possible that adding organic matter in a solid form (BSP) to the water may have caused the algal growth to accelerate, meaning that it used up the nitrates faster. The BSP also had a generally lower transmittance than the other two treatments. One possible explanation for this is the fact that as the BSP dissolved, the particles seeped out of the mesh bag, giving the water a darker appearance.

In general, the results varied greatly from sample to sample, and neither of the treatments appeared to be effective at inhibiting the algal growth when compared to the control. For example, one bucket containing BSE appeared clear, while a bucket right next to it containing the same treatment appeared green and cloudy. The variations in the results may have been due to random inoculation with different species of algae. Studies have shown that barley straw may be species-specific with inhibition of algal growth, which may explain why neither the extract nor the pellets were significantly effective (2). Another point to consider is that although the samples were all filled from the same water source and placed in the same greenhouse, each sample was still slightly different, containing its own unique array of organisms and nutrients. It is extremely difficult to understand precisely what is occurring in each sample. If I were to do this project again, I would test a larger sample size, and I would inoculate the sample with a specific strain of algae to test the effectiveness of the treatments on that specific strain. This would require me to do the tests in a lab environment. However, I do still see the value in conducting tests with random inoculation because in a realistic environment, there is no control over which species of algae begin to grow in the water. In addition, I would like to test phosphorus levels in water because research indicates that algae feeds on phosphorus, and may, therefore, be an indicator of potential algal blooms (13).

MATERIALS AND METHODS

Setup

We set up a seven-foot by eleven-foot greenhouse in a backyard in the middle of August, which is the peak season for algal blooms (1). Using the guidelines given on the back of the bottle, we calculated the correct concentration of *CrystalClear* barley extract at 0.06 mL of extract per 3.5 L of water. We calculated the correct concentration of *CrystalClear Nature's Choice* barley straw pellets at 0.4 g of pellets per 3.5 L of water using the guidelines given on the container. Then, we filled eighteen 5 L buckets with 3.5 L of well water. We measured BSE using a micropipette and added it to six buckets. We placed BSP in small mesh pouches and added them to six buckets. We used the remaining six buckets as control samples. In a refrigerator, we kept a jug of well water,

as well as a jug of the BSE solution, to be used as blank samples. We numbered the buckets and randomly assigned them spots in the greenhouse (**Figure 1**). Then, we placed mosquito netting over the buckets to prevent large insects from contaminating the samples. We conducted tests over a course of 25 days, with each data collection occurring approximately every five days. We allowed algae to grow in the buckets naturally through the spores in the air.

Data Collection

At each data collection, we scrubbed and stirred each sample with a separate wire brush to distribute the algae. We collected samples from the surface of each bucket in 5 mL test tubes using a pipette. On days when we collected hemocytometer tests, we collected additional samples in small vials. We took pictures of the samples each day that tests were conducted.

Strip tests

To conduct water quality tests, we dipped JNW Direct aquarium test strips into each bucket and compared the coloration of the indicators to the parameters given on the test bottle. The same person recorded the data each time to ensure that the readings were made consistently. We recorded data for free chlorine, nitrate, nitrite, carbonate, total alkalinity and pH. We conducted these tests on days 1, 6, 11, 15, 20, and 25.

DO Testing

We measured the DO levels by swishing a DO meter around in the water samples until the reading stabilized; then we recorded the data. We conducted these tests on days 1, 6, 15, and 25.

Spectrophotometer Testing

We measured the percentage of light transmittance using a spectrophotometer set at 700 nm wavelength. We collected samples in 5 mL test tubes and transferred them to 5 mL cuvettes. Before we tested each sample, we calibrated the machine using a blank sample (solution that had been kept in the refrigerator) and set it to 100% light transmittance. For the BSE samples, this was a sample of the BSE solutwithout any algae in it. For the control and the barley straw pellet samples, this was a sample of well water without any algae in it. Once the machine was calibrated, we inserted the test sample and recorded the transmittance.

Once we tested all of the samples, we rinsed the cuvettes once in hot, soapy water and twice in hot water. We wiped any excess drops of water with a Kimwipe and set the cuvettes upside-down and left them to air dry. We took care not to get any fingerprints or scratches on the cuvettes, as they can make the results less accurate. We conducted these tests on days 1, 6, 11, 15, 20, and 25.

Hemocytometer Testing

We measured cell density using disposable hemocytometer slides. To conduct the tests, we shook each sample well to distribute the algae; then we used a micropipette to load the hemocytometer slide with 10 µl of the sample. Next, we placed the slide under a microscope and brought it into focus. We counted algal cells in the four corner squares and the center square. We conducted these tests on days 6, 15, and 25.

Data Analysis

For all of the data that we collected, we ran a single-factor ANOVA test followed by a Tukey-Kramar test to determine whether the differences between the treatments were statistically significant. We also ran a descriptive statistics test to determine standard error. In addition, we conducted correlation analyses comparing the transmittance tests to each of the other tests. This showed how all of the indicators of water quality related to algal growth since the percentage of light transmittance was the main indicator of algal growth we analyzed. Comparing the cell density tests to the transmittance tests served as a way to ensure that the results were not affected by an error in experimentation since both cell density and transmittance are indicators of algal growth.

ACKNOWLEDGEMENTS

I would like to thank Dr. Brad Elder from Doane University and Mike Archer from the Nebraska Department of Environmental Quality for their assistance with my experimental design and data analysis. In addition, I would like to thank Mrs. Gavers, my information technology instructor, for assisting me with my figures and my mom for helping me gather data.

Received: April 10, 2020
Accepted: September 12, 2020
Published: October 06, 2020

REFERENCES

1. Schmale, David G., *et al*. "Perspectives on Harmful Algal Blooms (HABs) and the Cyberbiosecurity of Freshwater Systems." *Frontiers in Bioengineering and Biotechnology*, vol. 7, no. 128, 2019, pp. 1-7, doi:10.3389/fbioe.2019.00128.
2. Ferrier, M. D., *et al*. "The Effects of Barley Straw (Hordeum Vulgare) on the Growth of Freshwater Algae." *Bioresource Technology*, vol. 96, no. 16, 2005, pp. 1788-95, doi:10.1016/j.biortech.2005.01.021.
3. Prygiel, Emilie, *et al*. "Efficiency Evaluation of an Algistatic Treatment Based on Barley Straw in a Hypertrophic Pond." *Journal of Environmental Engineering and Landscape Management*, vol. 22, no. 1, 2014, pp. 1-13, doi:10.3846/16486897.2013.801847.
4. Trainer, Vera L., and F. Joan Hardy. "Integrative Monitoring of Marine and Freshwater Harmful Algae in Washington State for Public Health Protection." *Toxins*, vol. 7, no. 4, 2015, pp. 1206-34, doi:10.3390/toxins7041206.

5. Caffrey, J. M., and C. Monahan. "Filamentous algal control using barley straw." *Hydrobiologia*, vol. 415, no. O, 1999, pp. 315-18, doi:10.1023/A:1003884211027.
6. Lembi, Carole A. "Barley Straw for Algae Control." *Aquatic Plant Management (2002): 1-8. EBSCOhost*. Web. 2017, ucanr.edu/sites/csnce/files/57540.pdf.
7. Mecina, Gustavo Franciscatti, *et al*. "Response of Microcystis aeruginosa BCCUSP 232 to barley (*Hordeum vulgare L*.) straw degradation extract and fractions." *Science of the Total Environment*, vols. 599-600, 2017, pp. 1837-47, doi:10.1016/j.scitotenv.2017.05.156.
8. Williamson, Ian. "Living Bath." *New Scientist*, vol. 186, no. 2501, 2005, pp. 21-21, https://www.newscientist.com/lastword/mg18625012-800-living-bath/.
9. Franklin, P. A. "Dissolved oxygen criteria for freshwater fish in New Zealand: a revised approach." *New Zealand Journal of Marine and Freshwater Research*, vol. 48, no. 1, 2013, pp. 112-26, doi:10.1080/00288330.2013.827123.
10. "Understanding the Science of Ocean and Coastal Acidification." *United States Environmental Protection Agency*, www.epa.gov/ocean-acidification/understanding-science-ocean-and-coastal-acidification. Accessed 29 Aug. 2020.
11. Eddy, F. B., and E. M. Williams. "Nitrite and Freshwater Fish." *Chemistry and Ecology*, vol. 3, no. 1, 1987, pp. 1-38, doi:10.1080/02757548708070832.
12. "Free Chlorine Testing." *Centers for Disease Control and Precention*, 2014, www.cdc.gov/safewater/chlorine-residual-testing.html. Accessed 29 Aug. 2020.
13. Peeters, Edwin THM, *et al*. "Competition between Free-Floating Plants Is Strongly Driven by Previously Experienced Phosphorus Concentrations in the Water Column." *PLOS ONE*, 2016, pp. 1-18, doi:10.1371/journal.pone.0162780.

JOURNAL OF
EMERGING INVESTIGATORS

Article

Machine Learning Algorithm Using Logistic Regression and an Artificial Neural Network (ANN) for Early Stage Detection of Parkinson's Disease

Shreyas Kar[1], Peter W. Campbell[2]
[1] duPont Manual High School, Louisville, Kentucky, 40208, USA.
[2] W University of Louisville, Department of Anatomical Sciences and Neurobiology, Louisville, Kentucky, 40202,

SUMMARY

Parkinson's disease (PD) is the second most common neurodegenerative disease. Despite the prevalence of PD, diagnosing PD is expensive, requires specialized testing, and is often inaccurate. Moreover, diagnosis is often made late in the disease course when treatments are less effective. Since one of the earliest symptoms of PD are changes in voice patterns, we employed machine learning algorithms to detect these abnormalities. Using existing voice data from patients with PD and healthy controls, we created and trained two different algorithms: one using logistic regression and another employing an artificial neural network (ANN). The inputs for these algorithms were two statistical measures of voice patterns: Pitch Period Entropy (PPE) and Spread1. Both algorithms were successfully able to discriminate between PD patients and healthy controls with F2 scores > 0.93. Moreover, we found that the time from diagnosis had no impact on the performance of our models. Thus, we report the creation of two models that can reliably and accurately identify the voice patterns characteristic of PD. Our findings suggest that it is possible to diagnose PD by analyzing voice patterns, which would enable disease screening that is cheap, accessible, and accurate.

INTRODUCTION

Parkinson's disease (PD) is a neurodegenerative disease that is classically considered a motor disease but can impact the sensory and cognitive systems as well [1–7]. PD occurs when neurons in the brain die either via apoptosis (programmed cell death) or necrosis (unplanned cell death), which causes a decrease in dopamine, a neurotransmitter that is responsible for sending signals to other nerve cells, eventually leading to PD [8,9]. A number of factors contribute to the disease process including genetic predisposition, abnormal protein folding, oxidative stress, inflammation, and immune dysregulation [8,9]. Motor symptoms, especially tremors, usually present unilaterally with the left side involved more frequently than the right, but can spread bilaterally [10]. These symptoms arise because PD occurs due to a progressive degeneration of dopaminergic neurons [1,2,4–6,11]. Today the standard therapy is dopamine replacement therapy by administering the dopamine precursor, levodopa [3–5,12–14]. However, because PD is often detected at later stages when over 80% of dopamine neurons have died, current therapies, including levodopa, are often ineffective [7,14]. Dopamine neurons are located in the substantia nigra, a brain region which is involved in coordinating and planning body movement [1,6]. A reduction in dopamine from these neurons produces several motor symptoms, one of the earliest of which is difficulty in speech [2,11,15–17]. Voice symptoms are attributed to involvement of the vagus nerve and recurrent laryngeal nerve, which occurs either by pathological inclusions or by abnormal excitatory drive from the basal ganglia [18]. The voice of PD patients gets softer, breathy, slurred, and often mumbled. The tone of the voice becomes monotone without any inflection [7,17,19–23]. They also exhibit altered voice quality (dysphonia), a reduced range of articulation (hypokinetic articulation), and an irregular and rapid rate of speech (tachyphemia) [19]. In comparison to changes in brain imaging, PD associated voice changes are easier and cheaper to assess and occur earlier in the disease process. We therefore sought to create a method for detecting PD associated voice changes, which may be used to reliably, cheaply, and accurately diagnose PD. To diagnose PD using voice patterns, we took a machine learning approach that employed binary classification to assign inputs to one of two states: likely to have PD and unlikely to have PD. This classification can be accomplished by different types of algorithms[21,24–28], and we used two in this study: logistic regression and artificial neural networks. Logistic regression is a method of classification that finds the probability of a certain event occurring (in this case the probability of PD) as the output of the function [29]. An artificial neural network (ANN) is an algorithm, inspired by neuronal connectivity in the central nervous system, that takes in an input in the first layer, performs computations on the weighted inputs, applies activation functions in the hidden layer, and outputs a desired result in the final layer [30] . The connections between nodes and layers generates pattern recognition capabilities similar to those generated by neurons and synapses in animal brains.While ANNs are more powerful than logistic regression, they are also more costly in terms of computational demands [31]. Therefore, we used both methods to determine if there was a significant difference between the two approaches in this task. Others have taken similar approaches to use machine learning to assess PD [7,21–23,32,33]. However, these studies have relied on symptoms of PD that appear later in the course of the disease, such as gait disturbances and

tremor. While there have been some studies done on the early detection of PD, most have used algorithms such as support vector machines and data mining, not artificial neural networks [7,21,24–28]. As far as we are aware, this is the first attempt to use an artificial neural network to identify changes in voice characteristics to assist in the diagnosis of PD.

DISCUSSION

Here we report the development of two algorithms that were successfully able to discriminate the voice patterns of PD patients from healthy controls. Both the logistic regression and the artificial neural network (ANN) surpassed our success criteria (F2 > 0.87), and our analysis of the cost function indicates that the logistic regression was more accurate than the ANN. Moreover, the logistic regression had a specificity of 0.70 and a sensitivity of 0.95. Thus, while both algorithms were successful in identifying the voice patterns of PD patients, the ANN was more sensitive and had a higher F2 score.

We also investigated whether the years since diagnosis of PD impacted the function of our logistic regression. We found that the number of false positives and false negatives were similar to our original analysis yet the model poorly discriminated between early and late PD. The model performed similarly in discriminating normal from either early or late PD. This suggests that disease progression did not impact the performance of either model. However, due to how the data were obtained, we could not control for how long patients were in treatment or whether their treatments impacted our model [20]. Since our algorithms are intended to diagnose PD and not necessarily to be used to follow disease progression, the impact of levodopa, other pharmacotherapies, deep brain stimulation, or other therapies are beyond the scope of this study.

Our results support reports from others that machine learning is a promising approach to diagnose PD [7,21–28]. However, as far as we are aware, this is the first use of an ANN to examine voice patterns to diagnose PD. Some researchers have used advanced algorithms to analyze non-vocal data [31,32,39], and our success rate is similar. Others have examined vocal patterns using more traditional analyses [15,22,23,41], but logistic regressions and ANNs are generally preferable because of their ability for automated diagnosis, saving patients time and resources in early diagnosis. Thus, our data extend the field by employing machine learning algorithms to detect changes in vocal patterns.

This study demonstrates a relatively fast and inexpensive method for early stage detection of PD. Such a method can benefit patients as it may enable accurate diagnosis of PD at an earlier stage without expensive imaging [3,4,13,14,16,40]. Earlier diagnosis of PD may allow reevaluation of and improve the efficacy of therapies that have previously failed clinical trials due to use in late stage PD patients [11]. Furthermore, this method of detection can be administered remotely by using voice samples making it suitable for anybody around the world. This can be used as part of regular health checkup for populations susceptible to PD as this is both affordable and non-invasive. Once detected, a patient can then go through detailed clinical and physical diagnosis and start early treatment to reduce loss of valuable dopaminergic neurons.

MATERIALS AND METHODS

Data acquisition

Two voice parameters, Pitch Period Entropy and Spread 1, were used as inputs to the algorithms in this study. These data were obtained using the Machine Learning Repository of University of California Irvine [20]. This dataset is composed of voice samples of 195 voice recordings from 31 people. The voice recordings are comprised of 19 male voices and 12 female voices. The age varies from 46 to 81 with an average age of 66. Out of 31 people, 23 have PD. The number of years since diagnosis with PD varies from 0 to 28 among the PD patients.

Machine Learning Algorithms

Two major machine learning algorithms are used in this study: Logistic Regression and Artificial Neural network (ANN). A programing language called octave, similar to MATLAB, was used to implement these algorithms. Both algorithms were trained and tested using the acquired dataset. 60% of the data was randomly selected for training while the remaining 40% was reserved for testing. Spread 1 and PPE data were loaded into a matrix, termed X, and the status data was loaded into a column vector called Y.

Logistic Regression

The logistic regression was trained on 60% of the data using the flow chart illustrated in Figure 1. The cost function was implemented by using a vectorized version of the cost function for Logistic Regression. Next, in order to use an optimization algorithm, the partial derivatives with respect to each of the features was computed. The partial derivatives were stored in a column vector. After implementing the cost function and the gradients, the *fminunc* algorithm (finding minimum of unconstrained multivariable function) was trained over 100 epochs to find the optimal coefficient values. This algorithm numerically optimizes a multivariable function , in this case, the cost function of logistic regression. Optimal coefficient values (stored in a matrix called θ) are listed in the result section. The data were then normalized using two different techniques: z-score normalization and simple feature scaling normalization. After normalization of the training data, the algorithm was trained again. After logistic regression was implemented, the model was used to classify the test dataset. Prediction was performed by rounding the probability values. The decision boundary that separates the two classes is termed the hypothesis, which is determined by multivariate linear regression $y = \sum_{i=1}^{n} b_i x_i$ also written as $y = b^T x$).

Because probability values are always between 0 and 1, the output given by the hypothesis in multivariate linear regression needs to be constrained to output a value between 0 and 1. For this, the equation for multivariate linear regression is plugged into the logistic/sigmoid function, which is represented as:

$$g(z) = \frac{1}{1+e^{-z}}$$

If the value of $g(b^Tx)$ is greater than or equal 0.5 then the binary classification will output a value of 1, representing a high likelihood to get PD. However, if the value of the binary classification is less than 0.5 then the output will be 0, representing a low likelihood to get PD. The coefficients of the hypothesis are the optimal values such that the cost function is minimized. The cost function for logistic regression [34] is

$$J(\theta) = -\frac{1}{m}\sum_{i=1}^{m}[y^{(i)}\ln(h_\theta(x^{(i)})) + (1-y^{(i)})\ln(1-h_\theta(x^{(i)}))]$$

Where m represents the amount of training examples, y, in this case, represents the status of the Patient (0 for not having PD, and 1 for having PD). $h_\theta(x)$ represents the sigmoid function evaluated at a particular x value.

Artificial Neural Network (ANN)

Figure 1: Logistic Regression Approach

The cost function for the ANN was implemented using a *for* loop (representing the double summation) over all training examples classes (having PD and not having PD).

$$J(\theta) = -\frac{1}{m}\sum_{i=1}^{m}\sum_{k=1}^{K}[y_k^{(i)}\ln((h_\theta(x^{(i)}))_k) + (1-y_k^{(i)})\ln(1-(h_\theta(x^{(i)}))_k) + \frac{\lambda}{2m}\sum_{l=1}^{L-1}\sum_{i=1}^{s_l}\sum_{j=1}^{s_l+1}(\theta_{j,i}^{(l)})^2$$

Taking this sum and plugging it into the sigmoid function produced the value of each activation in the network. $z^{(n)}$ represented the activation values of the n^{th} layer before plugging into the sigmoid function, while $a^{(n)}$ represented the actual (post-sigmoid) activations of the n^{th} layer.

After the cost function was implemented, the backpropagation algorithm was implemented to compute the partial derivative with respect to each of the weights (θ_{ij}). This was done by first feedforwarding the neural network using a random set of weights (this was done for symmetry breaking). Then for each training example, an error value was computed for each node and accumulated in the *del one* and *del two* matrices. Finally, separate gradient vectors for $\theta(1)$and $\theta(2)$ were computed which were "unrolled" in a vector of partial derivatives, which was used in the optimization process along with the cost function.

As with logistic regression, the fminunc algorithm (with the gradient vector and cost function as inputs) was used to minimize the cost function and train the algorithm to compute the optimal values of the weights. The gradient vector was a column vector with the partial derivatives with respect to all the weight. It was calculated using a backpropagation algorithm and checked using Gradient checking (an algorithm that computes the partial derivatives numerically).

First, we performed forward propagation to determine the values of all the activations in the network. Next, all the "error" values, represented as δ, were calculated as

$$\delta = (\theta^{(m)})^T\delta^{m+1}.g'(z^{(m)})$$

except for the last output layer. That layer was determined as the value of the last activation layer subtracted by the value of the actual output.

Next, an accumulation term, Δ, was used to accumulate each of the errors. After this, it can be shown that $\frac{\partial}{\partial\theta_{ij}} = \frac{1}{m}\Delta_{ij}^{(l)}$ for the bias terms, and $\frac{\partial}{\partial\theta_{ij}} = \frac{1}{m}\Delta_{ij}^{(l)}\lambda\theta_{ij}^{(l)}$ for all other terms. While there does exist a simpler formula for $\frac{\partial}{\partial\theta_{ij}}$ when the regularization term is not accounted for (λ= 0) , this term helps prevent overfitting – a phenomenon which occurs when the model performes satisfactory with training data, but fails to generalize to test data – [35], which is of paramount importance in a neural network. This produced the gradient vector that was used as inputs to our optimization algorithm. After this the predict function was implemented by computing both activation values for each training example, and the largest one was the output.

Evaluation of Models

The predicted results are laid down against ground truth to evaluate each model so that True Positive (TP), True Negative (TN), False Positive (FP) and False Negative (FN) values are calculated **(Table 1)**. The general formula for F score is

$$F_\beta = (1+\beta^2) * \frac{precision * recall}{(\beta^2 * precision) + recall}$$

where precision is a function of TP and FP: precision = TP/(TP+FP); recall is a function of TP and FN: recall = TP/(TP+FN); and specificity is a function of TN and FP: specificity = TN/(TN+FP).

The algorithms were assessed based on their F2 scores. An F2 score is a weighted average of precision and recall and is used when minimizing the number of false negative cases is of greater importance than minimizing the number of false positive cases. The equation for F2 scores is as follows:

$$F_2 = 5 * \frac{precision * recall}{4 * precision + recall}$$

Algorithms were considered successful if F2 is greater than 0.87. This value was chosen because this algorithm is one of high recall, used when minimizing the amount of false-negative cases are of greater importance than minimizing the amount of false-positive cases [36–38]. Moreover, this value has been used by other researchers to evaluate model success [39].

Differentiating between PD stages

The same models were created for PD stage differentiation. The data set was split into 2 stages based on the 4-year mark. This specific benchmark was used as it is a common benchmark for early and late stage diagnosis

[40] . The split data set was evaluated based upon the earlier mentioned F2 score metric.

True Positive (TP)	Diagnosed PD vs Actual PD
False Positive (FP)	Diagnosed PD vs Healthy
True Negative (TN)	Diagnosed Healthy vs Actual Healthy
False Negative (FN)	Diagnosed Healthy vs Actual PD

Table 1: Describes the attribute of the confusion matrix

Normalization

Normalization is the process of converting data into a particular standard scale. This is done so that a data sample of a bigger range is not given additional preference solely due to its size. There exist three most commons forms of data normalization and standardization: Simple Feature scaling, z-scores and min-max normalization. Simple Feature scaling consist of converting the data set into values from 0 to 1. This is done of dividing each value in a column of a data set by the maximum value of the data set. If all the values are either strictly positive or negative then each example can be divided by the largest and smallest value respectively in the data set. *Z*-score normalization (more formally written as *z*-score standardization) is a method of normalization wherein each training example is converted into its respective *z*-score. This in turn converts the data set into one with a mean of 0 and a standard deviation of 1. This is done without changing the original shape distribution of the data set. The formula to calculate the z-score for a training example is given by the following equation:

Where z_i is the transformed value of the training example

$$z_i = \frac{x_i - \mu}{\sigma}$$

μ. epresents the mean of the data set and α represents the standard deviation of the data set.

Results

We constructed two different models to diagnose PD using two measures of voice patterns: Spread1 and PPE. These data were obtained from the Machine Learning Repository of University of California Irvine [20]. Both algorithms were trained on a random sample of 60% of the dataset and tested on the remaining 40%.

Logistic Regression

The first algorithm we trained was a logistic regression. We first normalized the data by converting raw values into respective z-scores and then dividing by the largest absolute value. This algorithm also uses three coefficient values that are stored in a matrix called Theta **(Table 2)**. Theta (1) has the constant value, Theta (2) has the coefficient values for Spread1 and Theta (3) has coefficient values for PPE. Using these values our logistic regression failed to pass the 0.87 threshold on the training data.

Therefore, we applied feature normalization to put Spread1 and PPE values in the same scale. This was done in two ways. Firstly, this was done by converting each number in a data set to its respective *z*-score , thus, normalizing the data set. In the second type of normalization, division, both data sets were divided by the highest absolute valued number in their respective data set. Then the absolute values of all numbers were taken after division. When the model was trained with the above normalized values, a different matrix for Theta was found **(Table 2)**.

	Initial Values	Z-Score Normalized	Divided Normalized
Theta 1	11.0235	2.43125	16.45383
Theta 2	1.7788	2.60668	16.97537
Theta 3	7.2889	-0.13596	-0.49524

Table 2: Theta Matrix for Logistic Regression

After these measures, prediction was performed with the test data and the algorithm achieved a F2 score of 0.93 **(Table 3).** Within the 78 voice patterns tested, the logistic regression achieved a specificity of 0.70 and a sensitivity of 0.95. The calculated F2 score of 0.93 was well above our threshold of 0.87 and indicates that this algorithm could successfully discriminate PD voice patterns. This is evidenced by our receiver-operator curve **(ROC; Figure 3),** which shows high specificity over various threshold values.

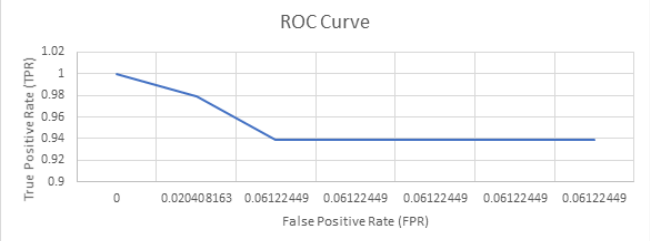

Figure 3: ROC Curve for Logistic Regression

Artificial Neural Network (ANN)

Initially, we implemented an ANN with one hidden layer. The ANN had two controllable matrices of weights (Theta), which were optimized using the fminunc algorithm and randomly initialized for symmetry breaking. After training, the model produced optimized weights **(Figure 2)**. The result of ANN with one hidden layer is in table 3.

The ANN with one hidden layer did not achieve the success criteria **(Table 3)**. To make the model more accurate and robust, we added an additional hidden layer. This layer was comprised of 4 activation units and a bias unit, and forward and backward propagation were performed to train the algorithm producing different optimized weights **(Figure 4)**.

Using the ANN with two hidden layers, our model passed the success criteria on the test data achieving an F2 score of 0.94 (table 3). The sensitivity of this model was 0.96,

the specificity was 0.61 using 78 voice patterns. The ROC curve **(Figure 5)** demonstrates high specificity over various threshold values.

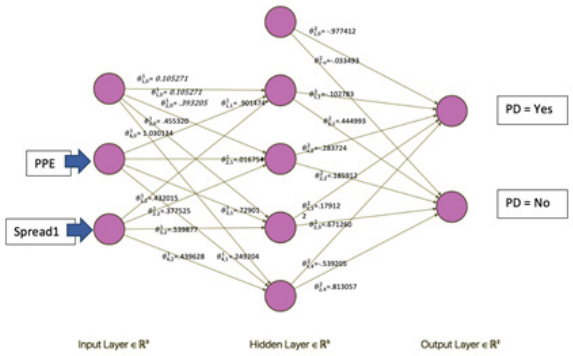

Figure 2: Weights and activations of ANN with single hidden layer

	Logistic Regression	ANN 1-Layer	ANN 2-Layers
Total Test Data	78	78	78
TP	54	53	53
FP	7	2	9
TN	14	12	14
FN	3	11	2
Precision	0.89	0.96	0.85
Recall/Sensitivity	0.95	0.83	0.96
Specificity	0.70	0.86	0.61
F1 Score	0.92	0.89	0.91
F2 Score	0.93	0.85	0.94

Table 3: Summary Data from Models

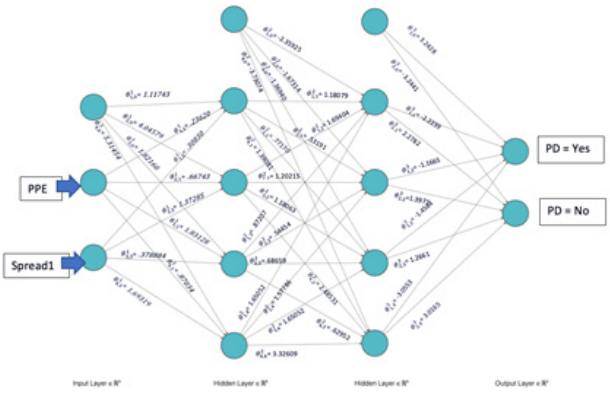

Figure 4: Weights and activations of ANN with two hidden layers

Differentiating between PD stages

PD is a neurodegenerative disease with multiple stages [2,4,5,7,12,14,16,36]. Therefore, we wanted to determine if the length of PD diagnosis had an impact on our model's performance. Patients with PD typically exhibit a response to L-DOPA therapy, as well as L-DOPA induced dyskinesias, within four years of diagnosis [40]. Moreover, the earliest that patients reached stage three of the disease was at 4 years [40], suggesting that a 4-year cutoff would differentiate early PD from late PD. Additionally, there are other neurodegenerative diseases, that while similar to PD in presentation, exhibit poor responses to L-DOPA (e.g. progressive supranuclear palsy, corticobasal degeneration) that are typically identified by this point in time [11]. Therefore, we divided patients into groups who had a PD diagnosis for less than 4 years or more than 4 years to determine if our results were confounded by disease stage.

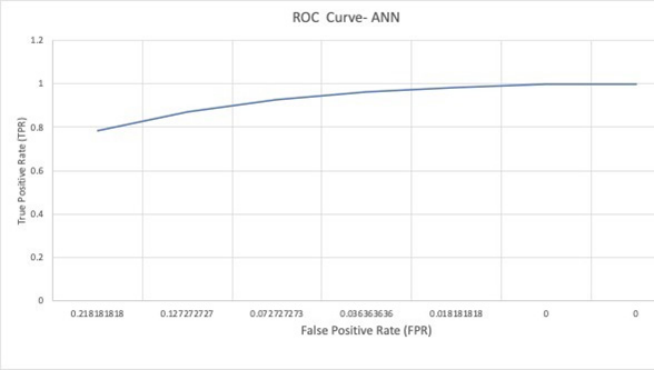

Figure 5: ROC Curve of ANN with two hidden layers.

We examined whether the logistic regression (used above) would perform differently on this divided dataset. The amount of true positive and false positive cases in the result for the test case when the algorithm was run was approximately the same. When the Logistic Regression was run to distinguish between the 65 PD patients who are below 4 years and above 4 years since diagnosis, with 49.2 % true positive, 47.7% false positive, 1.5% true negative and 1.5 %false negative were observed. This suggests that PPE and Spread1 are robust indicators and consistent indicators of PD and that our models were not impacted by the stage of PD. However, because there was little variability in these voice metrics, the logistic regression classifier performed poorly in identifying early PD from late PD (approximately 50% training accuracy, in this case). In other words, PPE and Spread1 values do not change significantly over time. Thus, this lack of variability means that the earlier stated algorithms can be used for detection at early stages.

Received: January 6, 2020
Acccepted: September 13, 2020
Published: October 10, 2020

REFERENCES

1. Wichmann, Thomas, and Mahlon R. DeLong. "Neurotransmitters and Disorders of the Basal Ganglia." Basic *Neurochemistry*, 2012, doi:10.1016/B978-0-12-374947-5.00049-3.
2. Kalia, Lorraine V., and Anthony E. Lang. "Parkinson's Disease." *The Lancet*, 2015, doi:10.1016/S0140-

6736(14)61393-3.
3. Chaudhuri, K. Ray, et al. "Non-Motor Symptoms of Parkinson's Disease: Diagnosis and Management." *Lancet* Neurology, 2006, doi:10.1016/S1474-4422(06)70373-8.
4. Pedrosa, David J., and Lars Timmermann. "Review: Management of Parkinson's Disease." *Neuropsychiatric Disease and Treatment,* 2013, doi:10.2147/NDT.S32302.
5. Olanow, C. Warren. "Levodopa: Effect on Cell Death and the Natural History of Parkinson's Disease." *Movement Disorders,* 2015, doi:10.1002/mds.26119.
6. Sonne, James, and Morris R. Beato. "Neuroanatomy, Substantia Nigra." *StatPearls,* 2018.
7. Bhat, Shreya, et al. "Parkinson's Disease: Cause Factors, Measurable Indicators, and Early Diagnosis." *Computers in Biology and Medicine*, 2018, doi:10.1016/j.compbiomed.2018.09.008.
8. Venderova, Katerina, and David S. Park. "Programmed Cell Death in Parkinson's Disease." *Cold Spring Harbor Perspectives in Medicine,* 2012, doi:10.1101/cshperspect.a009365.
9. Farooqui, Tahira, and Akhlaq A. Farooqui. "Lipid-Mediated Oxidative Stress and Inflammation in the Pathogenesis of Parkinson's Disease." *Parkinson's Disease,* 2011, doi:10.4061/2011/247467.
10. Heinrichs-Graham, Elizabeth, et al. "The Cortical Signature of Symptom Laterality in Parkinson's Disease." *NeuroImage*: Clinical, 2017, doi:10.1016/j.nicl.2017.02.010.
11. Ropper, Allan H., and Robert H. Brown. "Abnormalities of Movement and Posture Due to Disease of the Basal Ganglia." *Adams and Victor's Principles of Neurology,* 2005, doi:10.1036/0071469710.
12. Hornykiewicz, Oleh. "A Brief History of Levodopa." *Journal of Neurology*, 2010, doi:10.1007/s00415-010-5741-y.
13. Connolly, Barbara S., and Anthony E. Lang. "Pharmacological Treatment of Parkinson Disease: A Review." *JAMA - Journal of the American Medical Association,* 2014, doi:10.1001/jama.2014.3654.
14. Goldenberg, Marvin M. "Medical Management of Parkinson's Disease." *P and T,* 2008.
15. Trancoso, Isabel, et al. "Analysing Speech for Clinical Applications." *Lecture Notes in Computer Science* (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), 2018, doi:10.1007/978-3-030-00810-9_1.
16. Nutt, John G., and G. Frederick Wooten. "Diagnosis and Initial Management of Parkinson's Disease." *New England Journal of Medicine*, 2005, doi:10.1056/NEJMcp043908.
17. Goberman, Alexander M., and Michael Blomgren. "Parkinsonian Speech Disfluencies: Effects of L-Dopa-Related Fluctuations." *Journal of Fluency Disorders,* 2003, doi:10.1016/S0094-730X(03)00005-6.
18. Erman, Audrey B., et al. "Disorders of Cranial Nerves IX and X." *Seminars in Neurology,* 2009, doi:10.1055/s-0028-1124027.
19. Cannito, Michael P., et al. "Vocal Aging and Adductor Spasmodic Dysphonia: Response to Botulinum Toxin Injection." *Clinical Interventions in Aging,* 2008, doi:10.2147/cia.s1416.
20. Little, Max, et al. "Exploiting Nonlinear Recurrence and Fractal Scaling Properties for Voice Disorder Detection." *Nature Precedings*, 2007, doi:10.1038/npre.2007.326.1.
21.Belić, M., et al. "Artificial Intelligence for Assisting Diagnostics and Assessment of Parkinson's Disease—A Review." *Clinical Neurology and Neurosurgery,* 2019, doi:10.1016/j.clineuro.2019.105442.
22. Erdogdu Sakar, Betul, et al. "Analyzing the Effectiveness of Vocal Features in Early Telediagnosis of Parkinson's Disease." *PloS One*, 2017, doi:10.1371/journal.pone.0182428.
23. Lahmiri, Salim, et al. "Performance of Machine Learning Methods in Diagnosing Parkinson's Disease Based on Dysphonia Measures." *Biomedical Engineering Letters*, 2018, doi:10.1007/s13534-017-0051-2.
24.Dinov, Ivo D., et al. "Predictive Big Data Analytics: A Study of Parkinson's Disease Using Large, Complex, Heterogeneous, Incongruent, Multi-Source and Incomplete Observations." *PLoS ONE*, 2016, doi:10.1371/journal.pone.0157077.
25. Adams, Warwick R. "High-Accuracy Detection of Early Parkinson's Disease Using Multiple Characteristics of Finger Movement While Typing." *PLoS ONE,* 2017, doi:10.1371/journal.pone.0188226.
26. Bratić, Brankica, et al. "Machine Learning for Predicting Cognitive Diseases: Methods, Data Sources and Risk Factors." *Journal of Medical Systems*, 2018, doi:10.1007/s10916-018-1071-x.
27. Rovini, Erika, et al. "Comparative Motor Pre-Clinical Assessment in Parkinson's Disease Using Supervised Machine Learning Approaches." *Annals of Biomedical Engineering*, 2018, doi:10.1007/s10439-018-2104-9.
28. Cavallo, Filippo, et al. "Upper Limb Motor Pre-Clinical Assessment in Parkinson's Disease Using Machine Learning." *Parkinsonism and Related Disorders*, 2019, doi:10.1016/j.parkreldis.2019.02.028.
29. Bhattacharyya, Saptashwa. "'Logit' of Logistic Regression; Understanding the Fundamentals." *Medium - Towards Data Science*, 2018, towardsdatascience.com/logit-of-logistic-regression-understanding-the-fundamentals-f384152a33d1.
30. Chauhan, Nagesh Singh. "Introduction to Artificial Neural Networks (ANN)." *Medium - Towards Data Science,* 2019, towardsdatascience.com/introduction-to-artificial-neural-networks-ann-1aea15775ef9.
31. Parsaeian, M., et al. "Comparison of Logistic Regression and Artificial Neural Network in Low Back Pain Prediction: Second National Health Survey." *Iranian Journal of Public Health*, 2012.
32. Tsoulos, Ioannis G., et al. "Application of Machine Learning in a Parkinson's Disease Digital Biomarker Dataset Using Neural Network Construction (NNC) Methodology Discriminates Patient Motor Status." *Frontiers in ICT*, 2019, doi:10.3389/fict.2019.00010.
33. Tucker, Conrad, et al. "A Data Mining Methodology for Predicting Early Stage Parkinson's Disease Using

Non-Invasive, High-Dimensional Gait Sensor Data." *IIE Transactions on Healthcare Systems Engineering,* 2015, doi:10.1080/19488300.2015.1095256.
34. Ng, Andrew. "Machine Learning by Stanford University." *Coursera,* 2012, www.coursera.org/learn/machine-learning.
35. Nagpal, Anuja. "Over-Fitting and Regularization." *Medium - Towards Data Science,* 2017, https://towardsdatascience.com/over-fitting-and-regularization-64d16100f45c.
36. Powers, David M. W. "Evaluation: From Precision, Recall and F-Measure to ROC, Informedness, Markedness & Correlation." *Journal of Machine Learning Technology*, 2011.
37. Sasaki, Yutaka. "The Truth of the F-Measure." *Teach Tutor Mater*, May 2007.
38. Bzdok, Danilo, et al. "Statistics versus Machine Learning." *Nature Methods*, 2018, doi:10.1038/nmeth.4642.
39. Zhang, Kevin, and Dina Demner-Fushman. "Automated Classification of Eligibility Criteria in Clinical Trials to Facilitate Patient-Trial Matching for Specific Patient Populations." *Journal of the American Medical Informatics Association*, 2017, doi:10.1093/jamia/ocw176.
40.Poewe, Werner. "The Natural History of Parkinson's Disease." *Journal of Neurology*, 2006, doi:10.1007/s00415-006-7002-7.
41. Schwab, Patrick, and Walter Karlen. "PhoneMD: Learning to Diagnose Parkinson's Disease from Smartphone Data." *Proceedings of the AAAI Conference on Artificial Intelligence,* 2019, doi:10.1609/aaai.v33i01.33011118.

JOURNAL OF
EMERGING INVESTIGATORS

Article

Effect of Different Growth Media on Algae's Ability for Carbon Dioxide Biofixation

Shreya Chaudhuri[1] and Kara Pezzi[1]
[1] The Quarry Lane School, Dublin, California

SUMMARY

Carbon dioxide makes up 72% of all greenhouse gases produced, which makes it the leading source of air pollution. Certain green algal species such as *Chlorella vulgaris* fix the carbon dioxide into fatty acids present in cells in a process known as "carbon dioxide biofixation". This project tests the effect of different algal growth media on the efficiency of *Chlorella vulgaris*'s carbon dioxide biofixation. In the testing process, we added *Chlorella vulgaris* to four bottles, each containing four different substances (distilled water, Blue Green 11 medium, Bold's Basal Medium, and Guillard's f/2 medium), and cultured for eight days. Each algae and medium mixture was then divided equally into three smaller bottles and rotated for three days. To compare data, we measured the change in carbon dioxide content by subtracting the carbon dioxide content of the bottles with algae to a similar bottle without algae. The results for the average change in carbon dioxide content were 59.3 ppm for Blue Green 11 medium and algae, 50.6 ppm for Guillard's medium and algae, 22.6 ppm for Bold's Basal Medium and algae, and 10 ppm for distilled water and algae. The Blue Green 11 medium most effectively decreased carbon dioxide content of the bottles. This supported our hypothesis that algae's capacity for biofixation can be greatly enhanced through the effective use of media, a finding that has extensive real-world benefits in reducing pollution.

INTRODUCTION

Excess carbon dioxide in the atmosphere is one of the leading sources of air pollution. Carbon dioxide is a greenhouse gas, so when it is trapped in the atmosphere, the greenhouse effect causes a rise in global temperature, as the energy is not reflected back into space. There has been a 41% increase in carbon dioxide buildup in the atmosphere since 1990 due to excess air pollution. Carbon dioxide concentration has increased from 370 ppm in 1999 to just over 410 ppm in 2019, but 350 ppm is considered a safe concentration of carbon dioxide in the atmosphere (1). This excess carbon dioxide leads to ocean acidification, poor air quality, change in weather patterns, and a rise in sea levels. Air pollution is one of the world's biggest challenges to human health as well, as it kills up to nine million people annually and increases the risk of developing and aggravating lung diseases like asthma.

To avoid the devastating effects of air pollution, scientists have been researching feasible, sustainable options to reduce carbon dioxide pollution. Current research indicates that certain green algal species, such as *Chlorella vulgaris*, have the ability to sequester carbon dioxide into fatty acids present in their cells and use this carbon dioxide as a fertilizer to support more growth (2,3). Algae grow up to 50 times faster than terrestrial plants because they lack roots and shoots, which are energy sinks. Therefore, algal species have a higher photosynthetic efficiency than land plants, which have a photosynthetic efficiency of 1%. *Chlorella vulgaris* has a photosynthetic efficiency of 20% (4). Due to its high photosynthetic efficiency and growth rate, *Chlorella vulgaris* is commonly used in carbon sequestration research. Although this process of carbon dioxide biofixation is still under development, it has shown promising results of reducing air pollution up to 80%. Using algae to reduce air pollution is thus feasible and practical.

The goal of this research was to optimize the amount of carbon dioxide sequestered by the algae *Chlorella vulgaris* by altering the type of algal fertilizer or growth medium. We tested the effect of various growth media on the amount of carbon dioxide biofixation via *Chlorella vulgaris*. The growth media used in this project were Guillard's f/2 medium, BG 11 medium, and Bold's Basal Medium. All media used in this experiment are commonly used for growing freshwater algae like *Chlorella vulgaris* (5).

A higher quantity, mass, or cell density of *Chlorella vulgaris* results in higher displacement of carbon dioxide, as there would be more *Chlorella vulgaris* storing more carbon dioxide so the environmental carbon content would decrease (3). Consequentially, the medium which causes the most growth of algae would be directly responsible for the highest reduction in carbon dioxide.

Factors that affect the growth of *Chlorella vulgaris* include light, temperature, and pH levels. Light and temperature are kept at a constant, but different media have different pH levels. *Chlorella vulgaris* grows the best under media with higher pH levels, with the optimum pH for growth around 8.0 (5). Nutrients cause the differences in pH. Guillard's f/2 medium has a pH of approxiamately 7.5 to 7.9. BG 11 medium has a pH of 8.0. Bold's Basal Medium has a pH of 6.6. Distilled water (dH_2O) has a pH of 6.8, as after absorbing carbon dioxide from the air it becomes slightly acidic. Distilled water served

pH of media			
Medium	**Initial pH**	**pH after 7 days**	**pH after turning on the rotator**
BBM	6.65	7.28	7.36
BG 11	8.00	8.10	10.92
f/2	7.90	7.88	10.70
Distilled Water	6.90	7.50	6.53

Table 1: This table shows the change in pH levels over the entire experiment. As shown, the average pH of all of the mediums except the control distilled water is much lower initially than after turning on the rotator. The pH of the BG-11 medium in particular rises considerably post-rotation. It changes from a 8.0 to a 10.9. On the other hand, the control group's pH went down from 6.90 to 6.53.

as our control because it has no other minerals or nutrients. Therefore, I hypothesized that the medium with the highest pH will cause the highest growth in the algae and therefore, the highest reduction in carbon dioxide.

RESULTS

To test the hypothesis, I designed a procedure where I grew the algae in its appropriate medium under optimal circumstances, then transferred it to smaller, sealed bottles, and finally rotated it in these smaller bottles on a bottle rotator.

There were four stages to the experimental design. The first stage was building the culturing mechanism. The culturing mechanism consisted of four soda bottles with holes in the caps and tubing that connected the bottles to an air filter and air pump. The second step was making the growth media. I mixed in the appropriate amount of media with distilled water inside of each soda bottle and added *Chlorella vulgaris* for each media. For the control trial, I mixed the distilled water with the same amount of *Chlorella vulgaris*. The third stage was growing the algae in the culturing mechanism for eight days under optimum conditions. The final stage was to transfer the algae of each media into smaller sealed bottles and to rotate

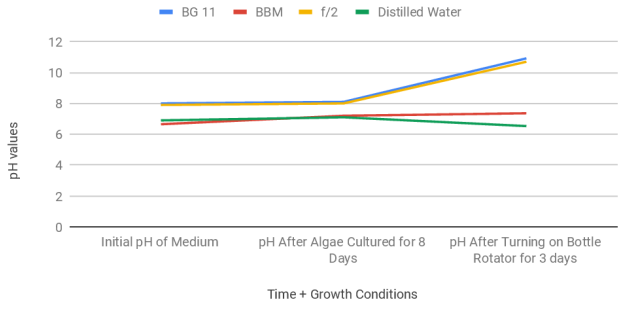

Figure 1: Effect of Time and Growth Conditions on pH Value of All Algae + Medium Mixture. This figure displays the correlation between the amount of time passing and the growth medium type on the pH value of the media and algae mixture. As seen in the graph, the pH increases for all of the trials except through the course of the experiment. The BG-11 medium that has the ratio that follows the Redfield Ratio the closest has the highest pH, while control groups' pH decreases during the experiment.

Average Cell Count of Each Medium + Chlorella vulgaris mixture	
Medium Name	**Average Approximate Total Cell Count (cells)**
BG 11	9.90×10^5
BBM	3.35×10^5
Guillard's f/2	8.25×10^5
Distilled Water	2.25×10^5

Table 2. This table shows the average cell count of each medium and *Chlorella vulgaris* mixture. The BG-11 medium had the highest cell count, and it was especially dramatic in comparison to the distilled water trial. The BBM's cell count was more similar to the distilled water trial, while the Guillard's f/2 medium's cell count was similar to the BG 11 medium's cell count. This pattern repeats with the other tables.

these bottles under the same conditions as before. The pH was measured by a pH meter and the carbon dioxide content was measured by a CO_2 gas sensor.

The cell count was measured with a hemocytometer after the bottles' rotation on the rotator. As portrayed in **Tables 1-4**, the BG 11 medium condition had the highest average cell density of 1320 cells per mL, the highest average cell count of 990,000 cells, the highest average change in carbon dioxide of 59.33 ppm, and the highest resulting pH level of 10.92. The BBM condition had an average cell density of 446 cells per mL, a cell count of 335,000 cells, an average change in carbon dioxide of 22.67 ppm, and a resulting pH level of 7.36. The Guillard's f/2 medium condition had an average cell density of 1100 cells per mL, a cell count of 825,000 cells, an average change in carbon dioxide of 50.67 ppm, and a resulting pH level of 10.7. The distilled water control condition had an average cell density of 200 cells/mL, a cell count of 225,000 cells, an average change in carbon dioxide of 10 ppm, and a resulting pH level of 6.53. Enough trials were not conducted for statistical analysis because lab was closed due to ongoing pandemic.

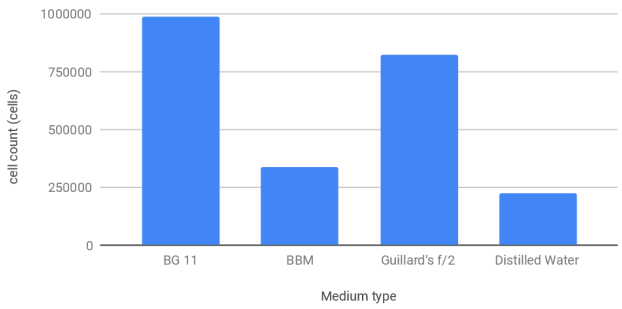

Figure 2: Average Cell Count of All Algae + Medium Mixtures (avg. of all trials) This figure displays the average cell count produced by each of the media and algae mixtures. The cell count is the highest for the BG-11 medium, which fosters the largest growth due to its nutrient levels. The control group has the lowest cell count due to the absence of nutrients. This pattern repeats for the **Figures 3 & 4**.

Average Cell Density of Each Medium	
Medium Name	**Average Cell Density (number of cells/mL of media)**
BG 11	1320
BBM	446
Guillard's f/2	1100
Distilled water	300

Table 3. This table depicts the average cell density of each medium and *Chlorella vulgaris* mixture. Again, the BG-11 medium has the highest average cell density, while the control had the lowest cell density. The BBM trials followed close to the control's cell density, and the Guillard's f/2 was similar to the BG-11's cell density.

Average Change in CO_2	
Medium Type	**Change in CO_2 (ppm)**
BG-11	59.33
BBM	22.67
Distilled Water	10.00
f/2	50.67

Table 4: This table portrays the average change in CO_2 content in the water bottles influenced by the medium and Chlorella vulgaris mixture. The BG-11 medium has the greatest change in CO_2, while the f/2 is similar but does not produce as great of a difference. The distilled water has the lowest change in CO_2, and the BBM follows close to this change.

DISCUSSION

The presence of carbon dioxide lowers the pH of media, as interaction between carbon dioxide and water creates carbonic acid. If the algae stores more carbon dioxide, then the carbon dioxide content of the water will decrease, so the pH of the water will increase. We used this as an indicator to see how effective each medium was at decreasing carbon dioxide content. If the pH was high after rotating on the bottle rotator, then the algae likely decreased the carbon dioxide content. We observed this result, particularly with the BG 11 medium, which had a pH of 10.97 in pH units, while the distilled water had a pH of 6.53 in pH units (**Figure 1**).

BG-11 media is more alkaline. An optimum pH for most algae is around 8.2-8.7 (5). The starting pH of the BG-11 culture medium was 8.00 in pH units; of all of the media, this pH was the closest to the optimum pH, which might have influenced more algal growth (6).

Both N (nitrogen) and P (phosphorus) boost the growth of *Chlorella vulgaris*. N and P can raise the pH of the solution. They are present in water in the form of ammonia, nitrates, and phosphates, usually sourced from fertilizer and pesticide runoff into rivers or other water bodies. It is important to note that cell viability was not tested; this is a limitation to this research that should be considered in this discussion. It would be interesting to consider and test this out in future tests. The results from this research as seen in **Figures 2 - 4** are suggestive but ultimately inconclusive because statistical analysis was unable to be conducted (due to too few trials and lab closure from ongoing pandemic), so a further study must be conducted to confirm the results.

In previous literature, the amount of carbon dioxide biofixation by *Chlorella vulgaris* in 72 hours was an average of 38 ppm (7). In this study, the *Chlorella vulgaris* grown in BG-11 had an average carbon dioxide biofixation amount of 59.33 ppm, which is about 20 ppm higher than the amount present in the literature search. Therefore, this research did provide information on how to increase carbon sequestration to reduce more carbon dioxide pollution.

All three medias used sodium nitrate as the source of nitrogen. The Guillard's f/2 medium contained 1 gram of sodium nitrate. BG 11 medium contained 1.5 grams of sodium nitrate. Bold's Basal Medium contained 0.25 grams of sodium nitrate. This concentration of sodium nitrate corresponded with both the algal growth and the change in carbon dioxide. The relatively high nitrogen content of the BG 11 medium aided photosynthesis and reproduction of *Chlorella vulgaris*, so the *Chlorella vulgaris* grew faster. BG 11 medium had the highest average cell count of $9.90x10^{5}$ cells and cell density of 1320 cells per mL and therefore the highest change in carbon dioxide.

A source of error in this project could be the variability of each cell of algae used in the project. Each *Chlorella vulgaris*

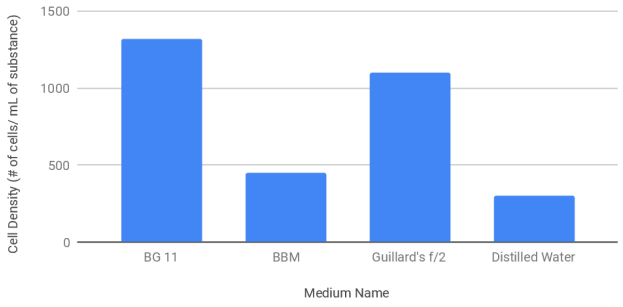

Figure 3: Average Cell Density of All Algae + Medium Mixture (avg. of all trials) This figure displays the average cell density of each of the media and algae mixtures. The cell density is highest for the BG-11 medium and lowest for the distilled water trials. The different is slightly less dramatic in comparison to the **Figure 2** but BG-11 still has the highest level of cell density and the control the lowest.

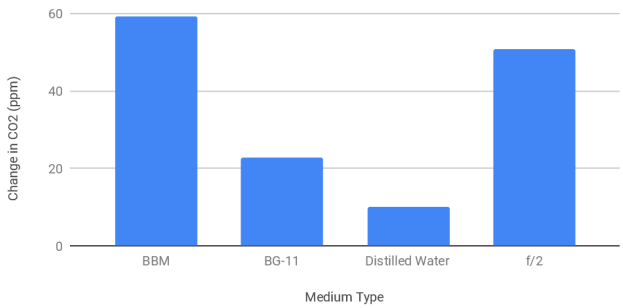

FIgure 4: Average Change in CO_2 Concentration of All Medium + Algae Mixture. This figure displays the average change in carbon dioxide by each of the different algae and media mixtures. This graph is what demonstrates the main objective of this study. The BG-11 medium produced the greatest change in CO_2 content, as it had the greatest cell count and densities, while the control led to the least change, having the lowest cell count and density.

Medium Name	Redfield Ratio (N:P)
BG 11	37.5:1
BBM	1:1
Guillard's f/2	1:1
Distilled water	N/A (no known nutrients added)

Table 5: This table shows the Redfield ratios of the different mediums. As seen, only the BG-11 medium is remotely close to the 16:1 ratio. The control group does not have any added nutrients, while the Guillard's f/2 and BBM have the same ratios.

cell is different so it could lead to differences in growth and therefore a lower or higher carbon dioxide biofixation. Additionally, sometimes the optimal conditions could change because it was difficult to make sure that the conditions were consistently the same; this could affect algal growth as well as the carbon dioxide biofixation. To maintain the conditions as close to optimal throughout the project, the light intensity and temperature were constantly checked.

Recent research indicates that nitrogen is a more effective algal fertilizer than phosphorus because nitrogen influences more reproduction (8,11). The Redfield Ratio, a ratio for nutrients present in various algae and in ocean chemistry, indicates a mass ratio of 16:1 for nitrogen:phosphorus for optimal growth, so nitrogen must be available in higher quantities compared to phosphorus (11). As shown in **Table 5**, BG 11 medium contained 1.5 g of sodium nitrate and 0.04 g of potassium phosphate, a ratio of 37.5:1. BBM contained 0.25 g of sodium nitrate and 0.25 g of potassium phosphate, a ratio of 1:1. The Guillard's f/2 medium contained 1.0 g of sodium nitrate and 0.74-1 g of monosodium phosphate, with a ratio of around 1:1. So, the BG 11 medium had a much higher ratio of nitrogen to phosphorus, and grew the fastest and reduced the most carbon dioxide content. The ratio was more than double of the Redfield ratio, but this was most likely due to the BG 11 being created for small aquariums or for a few batches of algae, not for the ocean (10).

My findings indicated that the BG 11 medium was the most effective at reducing carbon dioxide through *Chlorella vulgaris*. In three days, *Chlorella vulgaris* grown in BG 11 medium in a 250 cm3 volume was able to remove 59.3 ppm from the air (**Figure 4**). In one year, *Chlorella vulgaris* grown in BG 11 medium in a 250 cm3 volume may be able to remove almost 7200 ppm from the air if kept in ideal conditions. Extrapolating that finding, an algal farm with *Chlorella vulgaris* grown in BG 11 medium of the size of a football field with area 10800 meters at a depth of 3 meters can remove almost 2 million ppm from the air in one year. This is based purely on calculations, but it is an area to potentially explore.

Chlorella vulgaris grows better when supplied with more nitrogen and phosphorus, both of which are found in waste or sewage material (9). So, if algal farms are supplied waste material, the algae will serve a dual purpose of reducing air pollution as well as water pollution. Keeping the algae at optimum conditions will certainly lead to high maintenance costs, but it would be important to examine whether there is a trade-off between the maintenance costs and the reduction of air pollution to see whether this is a viable and practical solution to this ever-growing issue. Another area to explore in the future would be to find the optimum ratio of nitrogen: phosphorus that would be beneficial for reducing air pollution through carbon sequestration by algae.

METHODS

To make the culturing mechanism, I connected the following items in order: an air pump, 10 cm of 0.64-cm inside diameter (ID) piping, 25.4 cm of the 0.64 cm outside diameter (OD) piping, an inline air filter, 25.4 cm of the OD piping, an elbow hose fittings (screwing the other end of the barbed adaptor to the back of the 4 port manifold). I screwed four elbow hose fittings to the front of the four-port manifold and attached a 50.8 cm of the OD piping to each of the four barbed adaptors. I drilled 0.64 cm holes to the caps of 4 soda bottles, and pushed the opposite end of the 50.8 cm of the OD piping into the hole for each of the pipings attached to the 4 barbed adaptors on the front of the four-port manifold. (**Figure 5**).

Making the different growth media

To make the growth media, I added 0.529 g of the Bold Basal Powder mixture to 750 mL of distilled water. I added 1.0 mL of 0.1 M sulfuric acid solution, and then adjusted the pH to 6.6 with potassium hydroxide. Finally, I added 37.5 mL of *Chlorella vulgaris* algae solution to the media, closed the cap of the bottle, and pushed the 50.8 cm of 0.64 cm outside

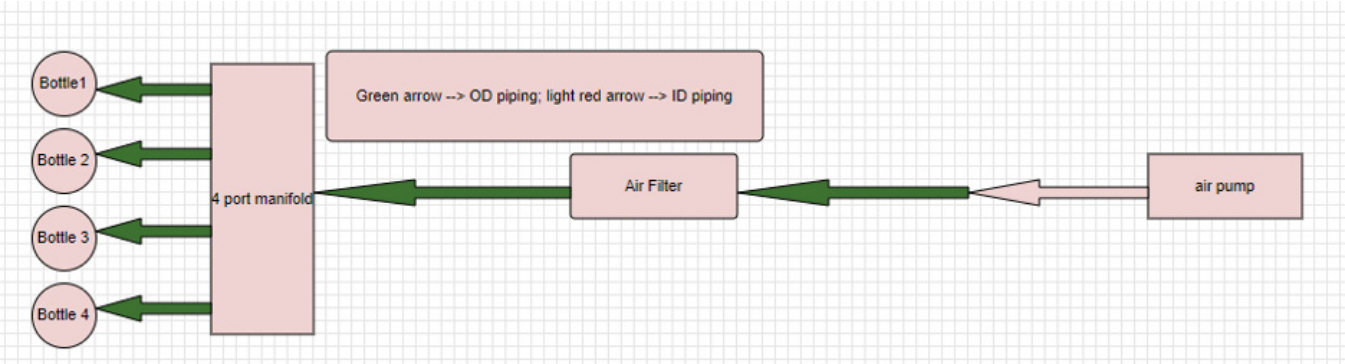

Figure 5: Schematic for assembling the culturing mechanism.

diameter tubing so that it was halfway in the middle of the media and algae mixture. Then, I added 1.26 g of the Blue Green 11 Powder mixture to 750 mL of distilled water. Next, I adjusted the pH to 8.0 with potassium hydroxide. I added 37.5 mL of *Chlorella vulgaris* algae solution to the media. I closed the cap of the bottle and pushed the 50.8 cm of the 0.64 cm outside diameter tubing so that it is halfway in the middle of the media and algae mixture.

I added 0.375 mL of the Guillard's f/2 solution to 750 mL of distilled water. I also added 37.5 mL of *Chlorella vulgaris* algae solution to the media, closed the cap of the bottle, and pushed the 50.8 cm of the 0.64 cm outside diameter tubing so that it is halfway in the middle of the media and algae mixture. Lastly, I added 750 mL of distilled water to the last soda bottle.

I also added 37.5 mL of *Chlorella vulgaris* algae solution to the media. I closed the cap of the bottle and pushed the 50.8 cm of the 0.64 cm outside diameter tubing so that it is halfway in the middle of the media and algae mixture.

Growing the algae

Then, I began the process of growing *Chlorella vulgaris*. I made sure that the temperature inside of the bottle was around 22˚C with a thermometer and that the light amount was approximately between 3000 to 3900 lux with a light meter. This is the optimum growth conditions for *Chlorella vulgaris* (4). Then, I grew the algae on the 16-hour light and 8-hour dark cycle and used the culturing mechanism system to aerate the algae for 8 days. While it grew, I checked every 5-6 hours to make sure that the light, temperature, and aeration is optimal at all times.

Rotating the algae

After the algae grew for 8 days, I measured the pH of each soda bottle with a pH meter. For each type of medium, I labelled three 500 mL water bottles with the trial number and medium name for a total of 12 bottles. Also, I made three more 500 mL bottles filled with 250 mL of distilled water but with no algae. Next, I mixed the algae in the soda bottles to evenly distribute the algae and added 250 mL of each medium including the control group into each of the three 500 mL bottles for the medium. For the control group for the water bottles, I added 250 mL of distilled water without any algae to the 500 mL water bottles. I put cling wrap around the opening of the bottles to avoid air release and close the lid and repeated this process for each of the media. I then fit the water bottles into the holes of the bottle rotator. I also made sure the environment has about 22˚C temperature and 3900 lux of light intensity and that the light dark cycle is 16-hour light and 8-hour dark. I grew the algae for three days on the bottle rotator.

After growing the algae for 3 days on the bottle rotator, I removed the water bottles and measured each of the water bottles' pH with a pH meter. Then, I calibrated the carbon dioxide gas sensor, set it at 0-100,000 ppm, inserted the sensor through the cling wrap of each bottle, and measured the carbon dioxide levels in ppm. I repeated this process with each water bottle.

ACKNOWLEDGMENTS

Special thanks to my biology teacher for helping me out with the cell density and count measurements, my robotics teacher for helping me build the bottle rotator, my chemistry teacher for helping me edit this article and measure the pH of the bottles, and both of my parents for financing this project and driving me to countless stores in search of materials.

Received: February 18, 2020
Accepted: October 2, 2020
Published: October 12, 2020

REFERENCES

1. "Carbon Dioxide Causes 80% of Global Warming." *Airclim*, www.airclim.org/acidnews/carbon-dioxide-causes-80-global-warming.
2. Adamczyk, Michał, *et al*. "CO_2 Biofixation and Growth Kinetics of *Chlorella vulgaris* and Nannochloropsis Gaditana." *Applied Biochemistry and Biotechnology*, Springer US, Aug. 2016, www.ncbi.nlm.nih.gov/pmc/articles/PMC4978769/.
3. Wang, B., and C.Q. Lan. "Biofixation of Carbon Dioxide (CO_2) by Microorganisms." *Developments and Innovation in Carbon Dioxide (CO_2) Capture and Storage Technology*, Woodhead Publishing, 27 Mar. 2014, www.sciencedirect.com/science/article/pii/B9781845697976500151.
4. Hanifzadeh, M. "Carbon Dioxide Biofixation and Biomass Production from Flue Gas of Power Plant Using Microalgae." *Carbon Dioxide Biofixation and Biomass Production from Flue Gas of Power Plant Using Microalgae - IEEE Conference Publication*, IEEE, 2012, ieeexplore.ieee.org/document/6190469/.
5. "Culturing Algae." *Carolina Biological Supply*, www.carolina.com/teacher-resources/Document/culturing-algae-instructions/tr29112.tr.
6. Rachlin, Joseph W. "The Effects of pH on the Growth of *Chlorella vulgaris* and Its Interactions with Cadmium Toxicity." *Archives of Environmental Contamination and Toxicology* link.springer.com/article/10.1007/BF01065839.
7. Shabani, M. CO_2 Bio-Sequestration by *Chlorella vulgaris* and *Spirulina Platensis* in Response to Different Levels of Salinity and CO_2. 2019, www.iaees.org/publications/journals/piaees/articles/2016-6(2)/CO_2-bio-sequestration-by-Chlorella-vulgaris.pdf.
8. Loreau, Michel. "Regulation of Redfield Ratios in the Deep Ocean." *AGU Journals*, John Wiley & Sons, Ltd, 27 Feb. 2015, agupubs.onlinelibrary.wiley.com/doi/full/10.1002/2014GB005066.
9. Durbin, E., and Chen. "Effect of PH on the Growth and Carbon Uptake of Marine Phytoplankton." *INT-RES*, www.int-res.com/articles/meps/109/m109p083.pdf.

10. Fu, Weiqi, *et al*. "Sugar-Stimulated CO_2 Sequestration by the Green Microalga *Chlorella vulgaris*." *Science of The Total Environment*, Elsevier, 9 Nov. 2018, www.sciencedirect.com/science/article/pii/S0048969718344759.
11. Loreau, Michel. "Regulation of Redfield Ratios in the Deep Ocean." *AGU Journals*, John Wiley & Sons, Ltd, 27 Feb. 2015, agupubs.onlinelibrary.wiley.com/doi/full/10.1002/2014GB005066.

JOURNAL OF
EMERGING INVESTIGATORS

Article

Transfer Learning for Small and Different Datasets: Fine-Tuning A Pre-Trained Model Affects Performance

Ananya Gupta[1] and Meghna Gupta[2]
[1] The International School Bangalore, Bangalore, Karnataka, India
[2] Adarsh Palm Retreat, Bangalore, Karnataka, India

SUMMARY

Machine learning and deep learning algorithms are rapidly becoming integrated into everyday life. Whether it is in your face-ID to unlock your phone or the detection of deadly diseases like melanoma, neural networks have been traditionally designed to work in isolation to achieve amazing tasks once thought impossible by computers. However, these algorithms are trained to be able to solve extremely specific tasks. Models have to be rebuilt from scratch once the source and target domains change and the required task changes. Transfer learning is defined as a field that leverages learnings and weights from one task for related tasks. This process is quite smooth if one has enough data and the task is similar to the previous, already learnt task. However, research on when these two conditions are not met is scarce. The purpose of this research is to investigate how fine-tuning a pre-trained image classification model will affect accuracy for a binary image classification task. Image classification is widely used, and when only a small dataset is available, transfer learning becomes an important asset. Convolutional neural networks and the VGG-16 model trained on Imagenet will be used. Through this study, I am investigating whether there are specific trends in how fine-tuning affects accuracy when used for a small dataset which is dissimilar from Imagenet. This will allow for the beginning of investigating quantifiable methods to train a model when using Transfer Learning techniques.

INTRODUCTION

The increasing use of artificial intelligence (computer systems that mimic human intelligence) (1), or AI, in myriad tasks and fields (2) has created a need for large amounts of data collection. Weights and biases are learnable parameters of a machine learning model which improve while the model trains on training data to provide an optimum model and correct predictions (3). Supervised learning is when the neural network is fed labelled data known to be correct (4). However, in this natural, chaotic world, the ability to collect large amounts of data and labelling data points is tedious, often restricted, and uses many resources which are not readily available. This could reduce one's ability to use supervised learning, one of the most common and powerful methods to train classification-based machine learning models. Supervised learning works by, for example, provide an image of a cat as input and the label "cat" as the output, and an image of a dog with the "dog" label. With this input and output, the network calculates loss based on its predictions from its weights. The network then uses an algorithm, to modify the weights and biases to better map the input to the correct output and to minimize loss. Hence, it predicts whether an image is a "cat" or "dog". As AI and machine learning tasks become more common in the general public, the field must reduce the need for large data to create accurate models. This would promote innovation that is truly democratic and open, free from the burden of having to amass large datasets. Doing so, AI is made less niche and more accessible for commercial and social use.

AI has been growing exponentially, getting closer and closer towards general AI which is AI that can learn to do myriad tasks, unlike specialized AI made for one specific task (1). One large step towards this has been the development of transfer learning. Transfer learning techniques are being used extensively due to their abilities to provide high accuracy by leveraging pre-learnt weights from a pretrained model to another (5). This field is significant because it reduces dependence on large datasets without compromising accuracy (6). Traditional machine learning techniques are based on the model of isolated, single task learning wherein knowledge from a past task is not leveraged for other tasks. Transfer learning, on the other hand, relies on previously learnt tasks which could allow the learning process to be faster, have higher accuracy and require less training (6). Transfer learning allows a source task to affect the inductive bias of the target task.

Fine-tuning is the process of unfreezing different layers in the base model (pre-trained model) (5). We refer to nomenclature in transfer learning of "unfreezing" which means "making a base layer trainable.''. Usually, transfer learning is conducted with deep neural networks by fine-tuning a pretrained model on the source task using data from the target task. This study investigates how fine-tuning a model (freezing and unfreezing different layers) will affect the performance (measured using accuracy) for small and dissimilar datasets. The VGG16 convolutional neural network, which was trained on the ImageNet dataset (6), is used as a base model for this investigation. A convolutional neural network is merely a deep learning algorithm which

Unfrozen Layer	5	4	3	2	1	0
Training Accuracy	**99.69%**	**99.69%**	**96.52%**	**91.99%**	**88.36%**	**87.76%**
Validation accuracy	**86.67%**	**90.00%**	**79.87%**	**76.67 %**	**73.47%**	**69.73%**

Table 1. Unfreezing lower layers causes validation accuracy to first increase and then decrease. It also causes training accuracy to stagnate and then decrease. These values were based on analogous training conditions with the only difference being which layer was unfrozen. These values are after 100 epochs of training.

takes in an input image, assigns importance (learning weights and biases) to various aspects or objects in the image and is able to differentiate one from the other. The target task is to classify cartoon illustrations of textbooks as male or female, which is quite different from the source task of the base model which is classifying objects (10). The dataset of 640 training images is also dissimilar to the Imagenet dataset, which is an important factor in this study. While other research has focused on how performance varies with fine-tuning, this investigation focuses on specifically small and different datasets (11-12). The size of training data used here is neither at the extreme of one-shot learning or the large datasets in deep neural networks. Hence, through these experiments we try to investigate whether quantifiable trends are noticed in how performance varies with fine-tuning patterns. We predict that quantifiable patterns will be noticed, as performance will first increase and then gradually decrease.

RESULTS

The pre-trained model used in this study is the VGG16 convolutional neural network, which was trained on the ImageNet dataset (6). On top of the base VGG16 model, further convolutional neural network layers are built. This specific experimental setup was chosen due to the widespread use of VGG16 in the Transfer Learning corpus, thus becoming a widely accepted model. By fine-tuning different layers and training with this small, dissimilar dataset we have found that the validation accuracy follows a clear trend with layers becoming unfrozen (**Table 1**). Validation accuracy first increases from 86.67% when layer 5 is unfrozen to 90.00% as the lower layer, layer 4 is unfrozen. This trend then shifts as validation accuracy continues decreasing when lower layers are unfrozen. There is a linear trend observed once layer 3 is reached, with validation accuracy falling linearly up to layer 0 being unfrozen. Hence, there is an initial increase and then a gradual fall in validation accuracy (**Figure 1**).

The training accuracy follows a distinct trend as layers become unfrozen (**Table 1**). Training accuracy first stagnates for layers 5 and 4 unfrozen, with an accuracy of 99.69%. After this plateau, there is a linear decrease in the training accuracy until layer 1. This is a relatively steep decrease. After reaching layer 1, there is a fall in the rate of decrease of the training accuracy. This means that the decrease per layer becomes lesser at the last data point, between layer 1 and 0. Hence, the graph follows the phases- stagnation, steep and linear decrease, and finally a gradual linear decrease. These phases are attained as lower and lower layers are unfrozen (**Figure 1**).

DISCUSSION

The trend followed by the validation accuracy and training accuracy is quite similar as seen in (**Figure 2**). The similar bell-shaped trends for both portray that measures of performance change in similar patterns as lower layers are unfrozen. This corresponds to the weights and biases learnt, allowing for a justified analysis since it conveys that the model is not overfitting as the validation and training accuracy have a similar trend. Overfitting is when the model only works well on a particular dataset (such as its train dataset) but is unable to perform well on other datasets (such as the validation or test dataset) (7). Since our validation and training accuracies

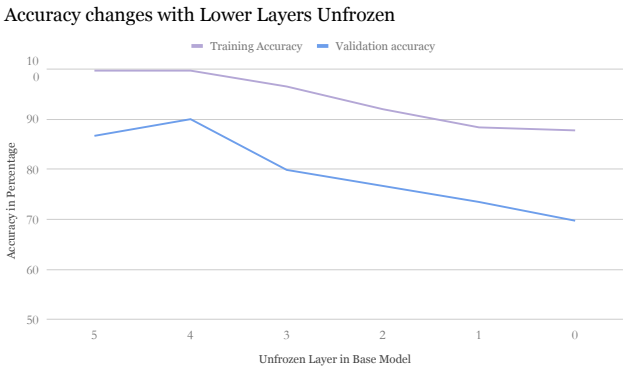

Figure 1. Unfreezing lower layers causes validation accuracy to first increase and then decrease. It also causes training accuracy to stagnate and then decrease. A line of the data of **Table 1** was plotted using linear interpolation. This interpolated line between each consecutive data point is used only to analyze the rate of change of accuracy with lower layers unfrozen for the discrete data values.

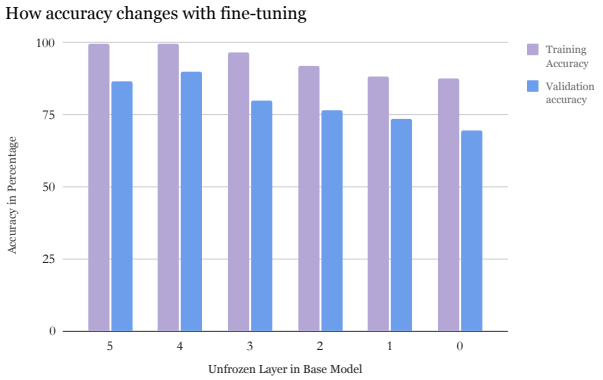

Figure 2. A bar graph showing how unfreezing lower layers causes validation accuracy to first increase and then decrease. It also causes training accuracy to stagnate and then decrease. Results of **Table 1** are used. Both types of accuracies were plotted to enable comparisons, portraying that the model is not overfitting.

are similar, it is fair to say that the model is not overfitting. Alongside, we note that the dataset used is relatively small, at only 640 training images, and is very different from the original Imagenet dataset. This serves the purpose of this study to investigate small datasets that are different from the original dataset. Data pre-processing is conducted in the form of Keras's data augmentation techniques which enables greater diversity of data available for training models, improving generalizations.

When only layer 5 is being fine-tuned, it is known that the weights learnt by the pre-trained model are more for the specific target task and domain of binary classification of male and female cartoon caricatures. This specificity is why it has the highest training accuracy and a high validation accuracy. This conclusion that specific training is better for a small dataset with data dissimilar from the ImageNet data is explained by the fact that there are very specific differences in the source and target domains. What this means is that for an unfamiliar task for the pre-trained model, general features such as edges are common but specific differences such as hair outline are extremely different, which makes specific training a necessity. However, its validation accuracy is lesser than when only layer 4 is being fine-tuned. This is because there is a fine line between being specific and being extremely specific wherein the general features are learnt incorrectly or become counterproductive. When only layer 4 is being fine-tuned, a great balance between the specificity and generality is struck. When taking a deeper look into the function of Layer 4 and 5 in-terms of specificity of weight and feature learning, this distinction becomes lucid.

Layer 5 has the potential to have higher validation accuracy if there was a greater amount of data, as this would lead to better generalizations.

When only layer 3 is unfrozen, there is a marked fall in training accuracy and validation accuracy compared to the fall in performance between unfreezing layers 5 and layer 4. This steep and sudden fall is at the hands of a shift in learning to more general weights. The weights learnt are not specific to the target task, causing a fall in training accuracy and validation accuracy as the model is not well trained for this dissimilar dataset. Layer 2 and 1 follow a similar trend of falling accuracies as less specific features and weights are learnt. Accuracies could be increased if there was a larger, more similar dataset, but for a small and dissimilar dataset this trend is followed while fine-tuning lower layers.

Finally, if no layers are fine-tuned, training accuracy and validation accuracy decreases once again. This is because the dataset is dissimilar to the ImageNet dataset. Hence, due to dissimilarities, the pre-trained model trained on ImageNet would not perform well on the current task. There are no updates to the weights, so the model is not learning new weights. Thus, it performs poorly for this dissimilar dataset.

Therefore, the results support the hypothesized trend that performance will first increase and then gradually decrease, as lower layers are left unfrozen. These results and corresponding theoretical justifications support the original hypothesis. With these results and hypothesis, we have found a quantified trend in how performance varies with fine-tuning as lower layers are unfrozen. These results on optimizing the fine-tuning combination to obtain higher model performance will allow better transfer learning models to be created for small, dissimilar datasets. One could further expand the scope of this investigation by exploring how performance changes with different combinations of unfrozen layers (specifically for VGG16, one could explore 32 combinations of 5C5 which gives 32). One could also perform similar investigations using different pre-trained models such as Inception, ResNet, VGG19 and so on to obtain the optimum fine-tuning combinations for these base models since this investigation focuses on the VGG16 base model.

MATERIALS AND METHODS

The VGG16 convolutional neural network pre-trained model, which was trained on ImageNet, was used for this experiment. VGG16 is a convolutional neural network model that consists of 16 weight layers, 13 convolutional layers, and 3 fully connected layers with 138 million total number of parameters (8). The model achieved 92.7% test accuracy in ImageNet (8). It has slowly become one of the most popular pre-trained models used, especially in Transfer Learning tasks, alongside models like ResNet, VGG19, and many more. It was trained on the ImageNet dataset which is a dataset of over 15 million labeled high-resolution images belonging to roughly 22,000 categories. The images were collected from the web and labeled by human labelers using Amazon's Mechanical Turk crowd-sourcing tool, ImageNet, which consists of variable-resolution images. Therefore, the ImageNet images were down sampled to a fixed resolution of 256×256. Given a rectangular image, the image is rescaled and cropped out the central 256×256 patch from the resulting

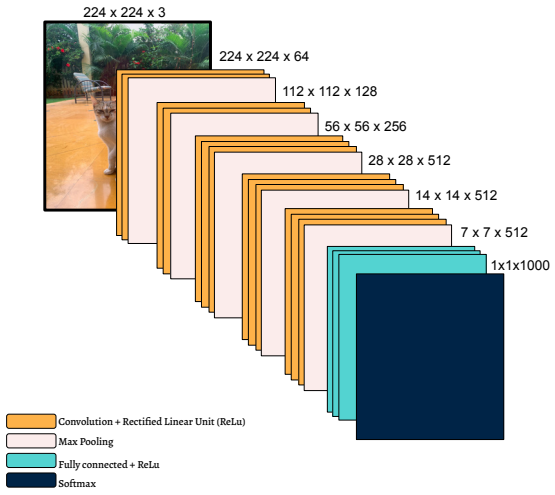

Figure 3. The VGG16 convolutional neural network pre-trained model architecture is detailed. The different types of layers, along with the input shapes are portrayed. A legend was also used to visualize how the model was built. (8)

image (9). The model architecture of VGG16 is depicted in (**Figure 3**) (6, 17).

Our dataset contains 2 classes - male and female textbook illustrations. It consists of 640 training images and 65 validation images. It was created manually since such a dataset does not exist to date, and alongside is dissimilar to the ImageNet dataset (which is a key part of this investigation) (10). We further confirmed that the images are dissimilar by attempting to classify the illustrations using purely the VGG16 model, which produced incorrect predictions. Alongside, after going through the categories classified in ImageNet, it is known that the dataset used in this investigation is different. We built the dataset by taking images from the Karnataka State Board textbooks in India, and our dataset can be seen here (10). Manual data collection was conducted to ensure uniqueness of the dataset. The data was split into 335 images of female illustrations and 337 male illustrations, to prevent any sampling bias. First, all the images are resized to 150x150 pixels, with 3 channels of red, green, and blue. The data is split into 3 datasets: training data, validation data, and test data. The images are converted to arrays, and each image is given a class of either male or female, i.e., 1 or 0. The images with their corresponding classes are shuffled. Shuffling helps reduce variance and reduce overfitting. It ensures that the gradient descent algorithm does not become "stuck" to a certain local minimum which prevents it from reaching the global minimum, by making sure that the input X changes with each iteration and is not static (13).

Data augmentation is effective when using small datasets in convolutional neural networks as it helps in reducing overfitting through rotation, rescaling, flip, etc. (14). For this project, the ImageDataGenerator factors of Keras are used for data augmentation and normalization of training dataset, which are detailed in **Table 2** below.

The ImageDataGenerator factors used for data augmentation and normalization of validation dataset are given by a Rescale factor of 1/255. These several parameters allow for powerful data augmentation techniques. The data augmentation techniques are applied on the data (**Figure 4**).

According to which layers are to be fine-tuned (or unfrozen), the model is built with setting, for example, block 5

Type	*Factor*
Rescale	1/255
Zoom Range	0.3
Rotation Range	50
Width Shift Range	0.2
Height Shift Range	0.2
Shear Range	0.2
Horizontal Flip	True
Fill Mode	Nearest

Table 2. The different data augmentation techniques used for the training dataset with their corresponding numerical factor. This is a part of our data pre-processing and is especially useful for small datasets like ours.

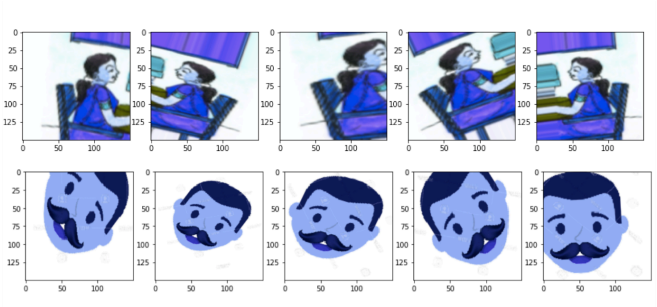

Figure 4. Sample of data augmentation conducted on the images in the dataset. The augmentation techniques are described in Table 2. Random combinations of data augmentation techniques were used. The image above displays examples of images from the two classes - 'male' and 'female'.

trainable as true. A Keras with tensorflow Sequential model is built, and the pre-trained model with the parameter of "trainable" set to "true" or "false" is added to the Sequential model. On top of that, convolutional neural network layers are added, as seen in **Table 3**. First, a Dense layer with a Rectified Linear Unit (ReLu) activation function is added. The Dense layer is a fully connected layer, which means that each neuron in a layer is connected to those in the following layer, and the parameters used here were 512 units (units are the output dimensions) and input dimension is the same as input shape. ReLu is a simple calculation that relays back the value given as the input directly or returns 0.0 if the number is ≤ 0.0. To train deep neural networks like the one used here and to use stochastic gradient descent with backpropagation of errors, an activation function is needed which can behave as if it were a linear function but in reality is a nonlinear function pertaining to complex relations in the data from which the model learns weights. The function must also allow greater sensitivity to the activation sum input and avoids easy saturation. Hence, the answer to all these criteria is the Relu activation function (15). It acts linear for all values greater than 0 but is actually a non-linear graph since it is flat for numbers ≤ 0.0. This is why an activation function is used for the Dense layer. Then, a Dropout layer is added with a 0.3 dropout rate. The function of the Dropout layer is to, at each step during training, randomly set input units to 0 with a frequency of 0.3. This is key to prevent overfitting and helps ensure better regularization. After this, another Dense layer followed by a Dropout layer is added, with the same parameters as specified before. The last layer added is a Dense layer with one unit to provide a one-dimensional output, along with the sigmoid activation function. The sigmoid activation function (also known as the logistic function) takes in the input and the function converts it into an output of a value in the range 0.0 to 1.0 (16). Binary cross-entropy is used to calculate loss since this is a binary classification problem, where y is the label and p(y) is the predicted probability of it being that class:

$$H_p(q) = -\frac{1}{N}\sum_{i=1}^{N} y_i \cdot log(p(y_i)) + (1 - y_i) \cdot log(1 - p(y_i))$$

It computes a score of the average difference between

the true and predicted probability distributions for predicting class 1 (17). Optimizers are algorithms which are implemented to vary the attributes of your neural network so as to obtain the minimum loss. They guide how a machine learning model learns and allows fast computation. They also prevent the model from believing that a local minimum has optimum weights because even though the loss function is a minimum at that local point, the optimum is only reached at the global minimum. This algorithm allows the model to find the globally optimum weights, serving its function as an optimizer. The RMSprop optimizer specifically does this by maintaining a moving average of square gradients and dividing the gradient by the root of this average, working on plain momentum, not Nesterov momentum. The effectiveness of this optimizer is why it is used for this task, and a learning rate hyperparameter of 1e-5 is also chosen. The learning rate was chosen such that the model does not get "stuck" at any local minima, and it also ensures that the model does not make any sudden, large updates to the weights such that it is unfavourable and harmful for the model (6).

Experiments were conducted on Google Colaboratory, with a 12GB NVIDIA Tesla K80 GPU. The model training is set with a batch size of 30 examples, 100 steps per epoch for 100 epochs and the Keras RMSProp optimizer. The validation steps were 50 steps. After 100 epochs, the training accuracy and validation accuracy are measured (**Table 1**).

Convolution2D	Conv Block 1
Convolution2D	
MaxPooling2D	
Convolution2D	Conv Block 2 - Frozen
Convolution2D	
MaxPooling2D	
Convolution2D	Conv Block 3 - Frozen
Convolution2D	
Convolution2D	
MaxPooling2D	
Convolution2D	Conv Block 4 - Trainable
Convolution2D	
Convolution2D	
MaxPooling2D	
Convolution2D	Conv Block 5 - Trainable
Convolution2D	
Convolution2D	
MaxPooling2D	
Flatten	VGG16 Base Model final layer
Dense	Added CNN layers
Dropout	
Dense	
Dropout	
Dense	

Table 3. The model architecture of the VGG16 convolution neural network. The layers within each of the blocks in the base VGG16 model, along with the added convolutional neural network layers are displayed. The base layers were optimized through the experiments.

Received: July 21, 2020
Accepted: October 13 2020
Published: October 18, 2020

REFERENCES

1. Wang, Pei. "On Defining Artificial Intelligence." *Journal of Artificial General Intelligence*, vol. 10, no. 2, 2019, pp. 1–37., doi:10.2478/jagi-2019-0002.
2. Poola, Indrasen. "How Artificial Intelligence in Impacting Real Life Every Day." *International Journal for Advance Research and Development*., vol. 2, Oct. 2017, pp. 96–100.
3. "Weight (Artificial Neural Network)." DeepAI, 17 May 2019, deepai.org/machine-learning-glossary-and-terms/weight-artificial-neural-network.
4. Amidi, Afshine, and Shervine Amidi. "Supervised Learning Cheatsheet Star." *CS 229 - Supervised Learning Cheatsheet, Stanford University*, stanford.edu/~shervine/teaching/cs-229/cheatsheet-supervised-learning.
5. Weiss, Karl, et al. "A Survey of Transfer Learning." *Journal of Big Data*, vol. 3, no. 1, 2016. Crossref, doi:10.1186/s40537-016-0043-6.
6. Shu, Mengying. "Deep Learning for Image Classification on Very Small Datasets Using Transfer Learning." *Iowa State University Digital Repository*, Iowa State University, 2019, lib.dr.iastate.edu/creativecomponents/345/.
7. Abu-Mostafa, Yaser. "DETERMINISTIC NOISE." Deterministic Noise - *Overfitting (Abu-Mostafa),* California Institute of Technology, 2012, work.caltech.edu/library/112.html.
8. Simonyan, Karen, and Andrew Zisserman. "Very Deep Convolutional Networks for Large-Scale Image Recognition." *ArXiv.org*, University of Oxford, 10 Apr. 2015, arxiv.org/abs/1409.1556.
9. Fei-Fei, L., et al. "ImageNet: Constructing a Large-Scale Image Database." *Journal of Vision*, vol. 9, no. 8, 2010, pp. 1037–1037., doi:10.1167/9.8.1037.
10. Gupta, Ananya. "Ananyagup/Classification-of-Males-and-Females-in-Textbook-Illustrations-." *GitHub*, 18 July 2020, github.com/ananyagup/Classification-of-Males-and-Females-in-Textbook-illustrations-.
11. Kandel, Ibrahem, and Mauro Castelli. "How Deeply to Fine-Tune a Convolutional Neural Network: A Case Study Using a Histopathology Dataset." *Applied Sciences*, vol. 10, no. 10, 2020, p. 3359., doi:10.3390/app10103359.
12. 12. Guo, Yunhui, et al. "SpotTune: Transfer Learning Through Adaptive Fine-Tuning." *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2019, doi:10.1109/cvpr.2019.00494.
13. Meng, Qi, et al. "Convergence Analysis of Distributed

Stochastic Gradient Descent with Shuffling." *Neurocomputing*, vol. 337, 2019, pp. 46–57., doi:10.1016/j.neucom.2019.01.037.
14. Wang, Jason, and Luis Perez. "The Effectiveness of Data Augmentation in Image Classification Using Deep Learning." *cs231n*, Stanford University, 2017, cs231n.stanford.edu/reports/2017/pdfs/300.pdf.
15. Agarap, Abien Fred. "Deep Learning Using Rectified Linear Units (ReLU)." *ArXiv.org*, Cornell University Arxiv, 7 Feb. 2019, arxiv.org/abs/1803.08375.
16. Jure Leskovec et al. "Neural Nets and Deep Learning." *Mining of Massive Datasets*, 3rd ed., Cambridge University Press, Cambridge, 2020, pp. 498–543.
17. Bishop, Christopher *M. Pattern Recognition and Machine Learning*. Springer New York, 2016.

JOURNAL OF
EMERGING INVESTIGATORS

Article

The effects of early probiotic supplementation on the germination of *Arabidopsis thaliana*

Julia Gambino[1], Mary Simons[1]
[1] Seaford High School, Seaford, New York USA.

SUMMARY

The agricultural industry uses fertilizer to produce enough food for the world's robust and growing population. However, the use of fertilizers is associated with an increase in soil degradation, which will lead to a decrease in crop production within the next decade. This soil degradation is due, at least in part, to the long-term use of oil-based synthetic fertilizers. These fertilizers deplete the soil of nutrients and bacteria essential to plant growth. It is critical to find solutions to support crop production to sustain the robust global population. As a result, this study was conducted to investigate how probiotic bacteria, like *Rhizobium leguminosarum*, *Bacillus subtilis* and *Pseudomonas fluorescens*, can impact the growth of *Arabidopsis thaliana* when applied to the seeds. The effect of multiple probiotic genera on early root growth was evaluated through the generation of various bacterial solutions composed of different quantities and combinations of probiotic bacteria, treating *Arabidopsis* seeds with these probiotic solutions and placing the seeds on agar plates to germinate with the evaluation of early root growth. We hypothesized that solutions with multiple bacterial species compared to those with only a single bacterial species would promote seed germination more effectively. The results suggest that there was no statistically significant difference in the total root lengths of each group, but there was a statistically significant difference in the number of root branches. Overall, treatment groups with probiotic species of bacteria promoted root branching more than treatment groups with no bacteria. Multiple species of these bacteria, especially the groups containing *R. leguminosarum*, stimulated more root branching than treatment groups with one species of bacteria. Further research on how different bacterial genera affect root characteristics may support the development of alternative and sustainable bacteria-based fertilizers, which can aid in reversing the effects of soil degradation.

INTRODUCTION

The pressure to produce enough food intensifies as the world's population continues to increase over time. The exponentially increasing global population is projected to surpass 9.8 billion people by 2050 (1). With more people comes the need for more food, and since agricultural commodities are the world's main source of nutrition, it is important that agricultural output keeps pace with the growing population. However, various environmental stressors, including the looming threat of climate change, have put the agricultural industry under immense strain. Another related environmental stressor is the impact of soil degradation on the agricultural industry (1). Farming an adequate amount of food becomes more difficult with the slew of deleterious effects resulting from soil degradation. These detrimental outcomes include the depletion of vital nutrients and the elimination of critical bacteria and fungi that synthesize essential organic material for plants from the soil (2). This soil degradation is due in part to the overuse of oil-based synthetic fertilizers that are typically used by commercial farmers to produce enough food for the robust global population. Utilizing these traditional synthetic fertilizers can cause severe soil degradation over time. This ultimately reduces the quality of the soil, decreases the amount of land viable for farming and impedes the ability of the soil to sustain plant life. About a third of the world's arable soil is degraded, and current degradation rates set the lifespan of conventional agriculture at a mere 60 years (3). This imminent future disaster calls for a change in the way we cultivate our food.

The use of biofertilizers is one of the most promising solutions for this impending crisis. Biofertilizers are live microorganisms that are added to the soil to enhance plant growth. This "living fertilizer" supplements the microbial community already present in the soil environment, referred to as the soil microbiome, instead of stripping the soil of its naturally occurring substances. A common example of microbes that enhance soil are growth-promoting rhizobacteria (PGPR) that assist plant growth when applied to the seeds or surfaces of a plant by colonizing the rhizosphere (4). These probiotic bacteria (bacteria with abilities that have positive health impacts) form a symbiotic relationship with the plant and survive while promoting plant growth. Using microorganisms as a form of fertilizer can boost crop production and enrich the nutritional quality of the products of commercial agriculture (5). In addition, biofertilizers aid in the restoration of the soil and other environmental damages associated with the use of these synthetic fertilizers. Unlike traditional fertilizers, biofertilizers are non-pollutant and inexpensive since they take in atmospheric nitrogen instead of depleting natural resources (6). This can be especially significant in poorer areas of the world where the need for a low-priced and effective way of maintaining crop production

is ever-present. Utilizing microorganisms to supplement plant and soil microbiomes in an effort to boost crop production is a sustainable and practical alternative to the synthetic oil-based fertilizers that jeopardize the Earth's ability to maintain the environment and our population. For example, a study that evaluated the effect of diverse microbial communities on plant growth found that more species of *Pseudomonas* helped plants grow more so than fewer species of *Pseudomonas* (7). Thus, further researching the potentials of PGPR can advance developments for sustainable forms of biofertilizer. These developments can help lessen the environmental damages that have occurred due to the overuse of traditional fertilizers. Although research has been conducted to evaluate the effectiveness of different probiotic bacteria, the full extent of the capabilities of these microorganisms remain unknown.

This led to us to ask: how can adding different amounts of probiotic bacteria from different genera to *Arabidopsis thaliana* seeds impact early root development when germinated on agar? We selected *Arabidopsis* for this experiment because, in addition to its fast growth cycle, there is an abundance of scientific knowledge pertaining to its specific genes and traits (8). This made *Arabidopsis* an ideal organism with which to conduct this experiment because procedures were carried out in a relatively timely manner, and the results obtained from this study can be modified and applied to other plants of greater agricultural significance. Additionally, studying *Arabidopsis* allowed us to examine the effects of microbes on studying seed germination, which is the first stage of a plant's growth cycle and is therefore vital for crop production and the establishment of a stable source of food for the world's population (9). Lastly, germinating the seeds on plain agar without any nutrients or outside factors successfully isolated the effects of the probiotic bacteria present. We utilized *Bacillus subtilis*, *Pseudomonas fluorescens*, and *Rhizobium leguminosarum* in this experiment, as they are three bacterial genera shown to have probiotic effects that positively impacted plant growth in previous studies (10, 11). We included two negative controls in this experiment. Our water-only group verified that the seeds could germinate on their own without the presence of bacteria for each trial. We included *Escherichia coli* as another negative control to distinguish the effects of a bacterium not known to be probiotic from the effects of bacteria known to have positive effects on plant growth like the species that made up the experimental groups. We organized different quantities and combinations of the three probiotic species selected into treatment groups to evaluate the effect of varying numbers of genera on plant growth. We hypothesized that a seed treated with a bacterial community composed of a larger number of bacterial species originating from multiple genera would have a greater impact on plant growth than communities made with a fewer number of species. The data collected from this experiment indicated that the number of probiotic bacterial species from different genera did not have a significant impact on total root length but more so on the number of root branches. Treatment groups with greater numbers of probiotic bacteria species promoted root branching more than treatment groups with one species. This signifies that the amount of probiotic species can play a substantial role in the development of root growth characteristics.

RESULTS

To determine the effect of multiple probiotic genera on early root growth, we conducted an experiment that involved the generation of different bacterial solutions to treat *Arabidopsis* seeds and monitor their growth. We grew overnight cultures of bacteria in nutrient-rich media, centrifuged the cultures to isolate the bacterial cells, and resuspended the cells in phosphate buffered solutions. We soaked sterilized seeds in combinations of bacterial solutions for 30 minutes and placed them on plain agar to germinate for two weeks. After, we measured total root length and the number of root branches with ImageJ software and analyzed these numbers using an ANOVA test.

We compiled abbreviations and descriptions for each treatment group **(Table 1).** "Bs" represents *B. subtilis,* "Pf" represents *P. fluorescens,* and "Rl" represents *R. leguminosarum*. Groups with multiple species of bacteria are denoted with these abbreviations connected with a plus sign, so "Bs + Pf" represents the group with *B. subtilis* and *P. fluorescens*, "Bs + Rl" represents the group with *B. subtilis* and *R. leguminosarum*, "Pf + Rl" represents the group with *P. fluorescens* and *R. leguminosarum*, and "Bs + Pf + Rl" represents the group with all three genera of bacteria. The negative control group with distilled water is portrayed as "water-only," and the negative control group with *E. coli* is named after that species.

The results obtained from this experiment can contribute to the knowledge needed to develop a viable alternative to traditional fertilizers and to making agriculture more sustainable for the sake of both environmental and global survival.

Water-only	*E. coli*	Bs	Pf	Rl	Bs + Pf	Bs+ Rl	Pf + Rl	Bs + Pf + Rl
Distilled Water	*E. coli*	*B. subtilis*	*P. fluorescens*	*R. legumino-sarum*	*B. subtilis* *P. fluorescens*	*B. subtilis* *R. legumino-sarum*	*P. fluorescens* *R. legumino-sarum*	*B. subtilis* *P. fluorescens* *R. legumino-sarum*

Table 1: The 7 experimental groups and the control groups that make up the different treatment groups are depicted in this table. Groups 1-7 are composed of different quantities and combinations of the three probiotic bacteria. The negative controls are the water-only and *E. coli* groups.

Total Length

We measured the total root lengths of the plants for each of the different treatment groups and calculated the average length per group. We found that the water-only and *E. coli* negative control groups had similar mean root lengths of 2.73 cm and 2.48 cm, respectively **(Figure 1)**. For the groups with treatments of only one species of bacteria (groups Bs, Pf, and Rl), the mean lengths were 2.54 cm, 3.52 cm, and 3.84

cm. Treatments consisting of combinations of two probiotic bacteria (groups Bs + Pf, Bs + Rl, and Pf + Rl) had mean lengths of 2.26 cm, 4.71, cm and 5.75 cm. Lastly, the treatment group featuring all three bacteria (group Bs + Pf + Rl) had a mean root length of 4.94 cm. The average root lengths of the experimental groups, apart from group Bs + Pf, were larger than those of either of the control groups. Additionally, groups Bs + Rl and Pf + Rl had larger mean lengths than groups Bs, Pf, and Rl. Group Bs + Pf + Rl had an average total length that was larger than every treatment group except for group Pf + Rl.

However, it is important to note that the data values for each treatment group were relatively inconsistent, as demonstrated by the ranges of root length **(Figure 1)**. After analyzing the total lengths with an ANOVA test, we determined that the differences in root length among the different treatment groups were not statistically significant ($P = 0.15$). Therefore, we could not draw definitive conclusions about the effectiveness of the quantity of probiotic bacteria from different genera on promoting plant growth based on total root length alone.

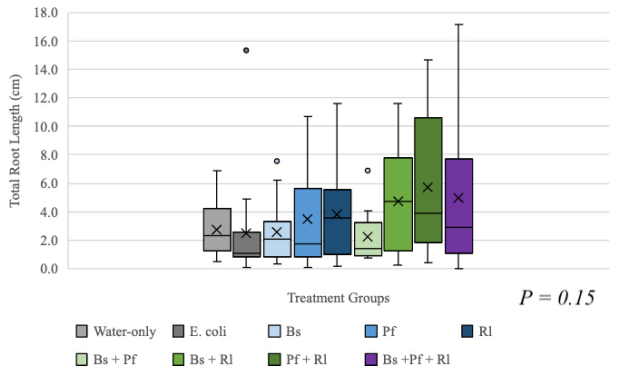

Figure 1: The total root lengths among treatment groups are depicted in this box and whisker plot. The "X" on each box denotes average values. The endpoint caps on each whisker indicate maximum and minimum values with the whiskers demonstrating the overall variability in each group. The lone dots represent outliers in the data set.

Number of Root Branches

We also counted the number of root branches per plant in each treatment group. The water-only and *E. coli* negative control groups had average branch counts of 0.714 and 0.526 branches, respectively **(Figure 2)**. Single species treatment groups (Bs, Pf, and Rl) had mean branch counts of 1.25, 2.50, and 3.13 branches. Treatment groups composed of combinations of two bacteria (Bs + Pf, Bs + Rl, and Pf + Rl) had average branch counts of 0.67, 2.87, and 3.11 branches. Finally, the treatment group consisting of all three probiotic bacteria (Bs + Pf + Rl) had an average branch count of 3.15 branches, which is the largest mean branch count out of all the treatment groups. The presence of probiotic bacteria led to higher numbers of root branches compared to the two control groups overall, but based on these values, there did not seem to be a distinct pattern between the average number of root branches and the number of bacterial species from different genera in the treatment groups. However, groups Rl, Bs + Rl, Pf + Rl, and Bs + Pf + Rl were the groups with the highest average branch counts and were the only groups to contain *R. leguminosarum*. Within these specific groups, the mean number of branches increased as the number of bacterial species in the treatment groups increased.

Just like the previous aspect of growth, the range of branch counts among the plants within each treatment group varied, as there was at least one plant with no branches in every group **(Figure 2)**. We analyzed the branch counts of each treatment group with an ANOVA test, which indicated that the differences in the average branch numbers among the different treatment groups were statistically significant ($P = 0.02$). Due to this, we concluded that both the presence and quantity of bacteria from different genera had an effect on the number of branches in this experiment.

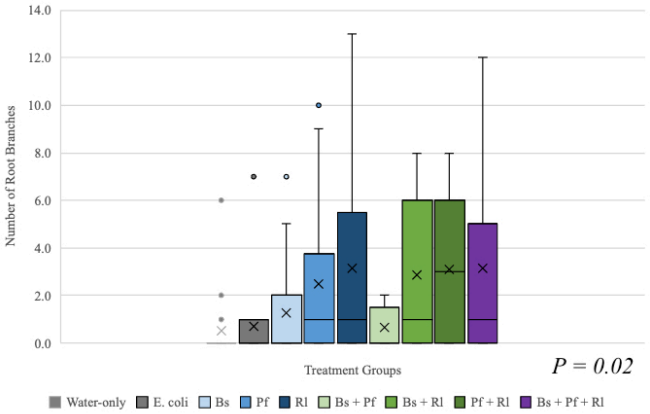

Figure 2: The number of root branches among treatment groups are depicted in this box and whisker plot. The "X" on each box denotes average values. The endpoint caps on each whisker indicate maximum and minimum values with the whiskers demonstrating the overall variability in each group. The lone dots represent outliers in the data set.

To further analyze the effectiveness of probiotic bacteria from multiple genera on root branching, we made multiple comparisons among the treatment groups. We performed a post-hoc Tukey HSD test to determine the statistical significance between specific treatment groups. This analysis concluded that between individual groups, there was no statistical significance among the results, as all the overriding *p*-values were greater than 0.05.

DISCUSSION

Based on our results, we rejected the hypothesis that a greater number of probiotic bacteria from multiple genera would lead to an increased promotion of root growth. The differences in total root lengths between the treatment groups were not statistically significant and could not be accurately utilized for further investigation, which led to the rejection of the original hypothesis. However, due to the overall

significant *p*-value for the number of root branches and the insignificant *p*-values comparing the number of root branches between groups, we chose to qualitatively analyze the results with pictures of the roots from the different groups. When examining specific plants in each of the treatment groups, it was clear that while the total lengths of the roots did not vary significantly, the roots of each plant in the different groups were distinct. An example of a plant from the water-only negative control group consisted of one continuous root that did not branch off in any direction **(Figure 3A).** In contrast, the main root of a plant from the group treated with all three genera of probiotic bacteria diverged into seven smaller roots that branched out into different directions **(Figure 3E)**. The roots of other treatment groups containing probiotic bacteria separated in a similar fashion **(Figures 3B, 3C, and 3D)**. When considering the different treatment groups and the specific species of probiotic bacteria that they were composed of, the diverging of the roots into the surrounding area of the agar plate appeared to be promoted by the presence of a greater number of probiotic species, as seen with the different groups containing *R. leguminosarum*. By qualitatively analyzing root characteristics in this way, we found that the quantity of probiotic bacteria from multiple genera may not significantly impact the growth of the roots with respect to total length but rather impact the way the roots grow and develop from germination.

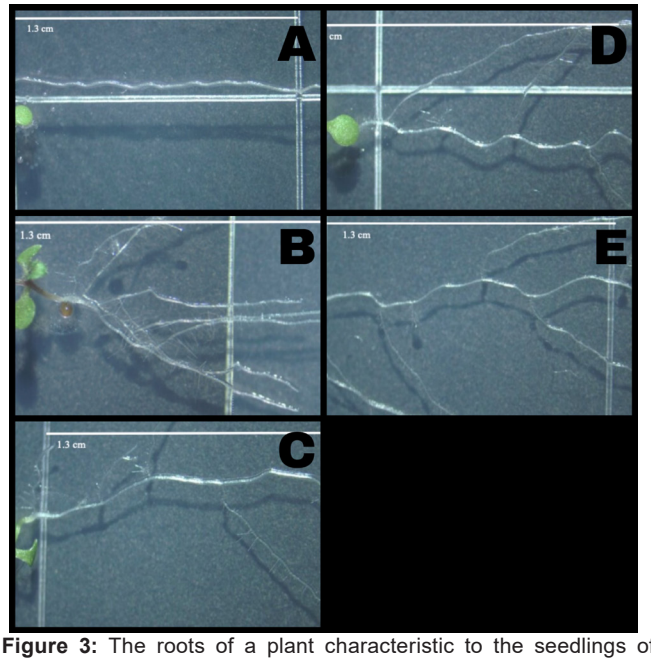

Figure 3: The roots of a plant characteristic to the seedlings of select treatment groups. A: distilled water; B: *R. leguminosarum*; C: *B. subtilis* and *R. leguminosarum*; D: *P. fluorescens* and *R. leguminosarum*; E: *B. subtilis*, *P. fluorescens*, and *R. leguminosarum*

Given that the original hypothesis was rejected overall, it is crucial to evaluate the various sources of error that were present in this experiment and the limitations that impacted the results. There was extensive variation in root length and the number of root branches between individual plants within the same treatment group, and there was also an unequal and inadequate sample size for each of the treatment groups. It was difficult to predict whether every seed would successfully germinate. During the two-week germination period, some of the seeds in each of the treatment groups failed to grow, and since we only analyzed the seeds that germinated for this study, the sample sizes for each group were unequal and relatively small by the conclusion of the data collection period. The small sample sizes of each group were not enough to counteract the variability present in the data sets, which led to a decrease in the validity of the results. Additionally, we could not capture the entirety of some of the plants with the camera-microscope setup due to the magnification settings of the microscope available. Therefore, when measuring the roots with ImageJ, we had to measure each picture separately. Since usage of the tracing tool stopped at the grid lines present in each picture, there is the possibility that there was some overlap from picture to picture, leading to slight errors in measurement. Lastly, although we carried out the process of cultivating the bacteria in the liquid media to create the probiotic solutions as efficiently as possible, we did not validate that the bacteria from every species survived in all the different treatment combinations. Thus, the validity of the results may have been impacted regarding the components of each treatment group.

Similar to how the probiotic bacteria in this experiment significantly impacted the number of root branches, further investigation of how probiotic bacteria can impact different root characteristics can be beneficial to increasing the knowledge in this field. The branching of roots into the space around the plant increases the amount of area that the roots extend into the soil. This allows the plants to both anchor themselves better into the soil for more support and gain nutrients from a greater area. These aspects portray the importance of further investigation of root characteristics like branching, root hairs, and the overall path of the roots and how that can be advantageous in increasing crop production and quality of growth. Additionally, since *R. leguminosarum* was associated with the four groups that had the greatest mean numbers of root branches, conducting similar studies with different probiotic bacteria from multiple genera that are related to *R. leguminosarum* could lead to significant developments. For example, *Rhizobium* is a genus that is categorized as an intracellular PGPR unlike *Pseudomonas* and *Bacillus,* which are both genera that are extracellular PGPR; *Bradyrhizobium*, *Sinorhizobium*, *Azorhizobium*, *Mesorhizobium*, and *Allorhizobium* are all bacterial genera that are considered to be intracellular PGPR and assist plant growth from inside the cells of the plants (12). Investigating the effects of *Rhizobium* bacteria in combination with bacteria from other intracellular PGPR genera that can work together from inside the cells of plants could be important and potentially increase the magnitude of the effect that *R. leguminosarum* had on the root branching in this study. Knowing more about how the specific development of roots can be manipulated by probiotic

bacteria from multiple genera can lead to advancements in the development of alternative and sustainable bacteria-based fertilizers, which can aid in reversing the effects of soil degradation. By supplementing the soil with probiotic bacteria, these bacteria-based fertilizers can help to alleviate the stress of the increasing global population on food production and enrich the agricultural industry. Continuing to find more uses for probiotic bacteria in this way can enhance the way we grow our food and extend the life span of agriculture for our planet.

MATERIALS AND METHODS

Cultivation of Bacteria

All bacteria cultures and the media and materials used to cultivate them were obtained from Carolina Biological Supply Company. We carried out proper laboratory precautions, such as wearing gloves, while preparing the bacteria and during any steps of this experiment in which these bacteria were utilized to ensure the safety of ourselves as the researchers. To prevent contamination that would affect the results of this experiment, we also created sterile conditions.

P. fluorescens, *B. subtilis* and *E. coli* were streaked on nutrient agar plates, and *R. leguminosarum* was streaked on *Rhizobium X* plates. We used a sterile plastic inoculation loop to pick up single colonies and spread them on their respective plates for cultivation. *B. subtilis* and *E. coli* were incubated at 30°C, and *P. fluorescens* and *R. leguminosarum* were incubated at 25°C. We streaked each species of bacteria on new plates weekly to keep a growing population of bacteria alive and prepared for when we were ready use them.

Proceeding this, we transferred each bacteria culture into a liquid medium. Nutrient broth was used for *P. fluorescens*, *B. subtilis*, and *E. coli,* and *Rhizobium X* broth was used for R. leguminosarum. The bacteria grew overnight on a rocking platform at a speed of 60 rpm. We then measured the optical density of each bacterial sample using a spectrophotometer at 600 nm with a 10 mL cuvette. After using a centrifuge to spin down and isolate the bacterial cells of each culture for 30 seconds at a speed of 6000 rpm and removing the remaining supernatant with a micropipette, we combined the cells with a phosphate-buffered saline (PBS) (13). The bacterial cell solutions were spun at the same speed for an additional 30 seconds and then suspended in PBS again.

These solutions served as the different treatments for the *Arabidopsis* seeds. We created these with either one, two, or three species of the probiotic bacteria in addition to the water-only and *E. coli* negative controls **(Table 1)**. For treatments with multiple species of bacteria, we took equal measurements of each solution to create a 500 µL mixture.

Seed Treatment

The *Arabidopsis* seeds were provided by Flinn Scientific, and we thoroughly sterilized them before experimentation. The seeds were rinsed with 500 µL of distilled water in a 1.5 mL tube by pipetting up and down, after which the distilled water was removed. This process was repeated with the same amount of a 5% sodium hypochlorite solution and again with 70% ethanol (14). Lastly, we rinsed the seeds with distilled water. We soaked the sterilized seeds in 500 µL of the different treatments for 30 minutes. After the seeds were treated, they were placed on clear square plates with sterile tweezers. Prior to seed placement, we filled the plates with a 0.8% agarose solution that set into a gel.

Germination and Growth

The square plates containing the seeds were placed on a rack at a 45° slant under a fluorescent light to germinate. We set the fluorescent light to be lit for a 16-hour photoperiod over a two-week span so that the seeds could fully germinate and develop into seedlings (15). The square plates used to germinate the seeds featured grid lines marking 1.3 cm, which we used as a scale to measure the root lengths of the different seedlings. We positioned a camera above a light microscope to take pictures of the seedlings. At the end of the two-week germination period, we took pictures of each seedling using this camera setup **(Figures 3A-E)**. The total number of seeds treated for each group was recorded in addition to how many of these seeds germinated by the end of the two-week period **(Table 2)**.

	Seeds Treated	Seeds That Germinated	Seeds That Did Not Germinate
Water-only	37	19	18
E. coli	19	14	5
Bs	24	16	8
Pf	24	16	8
Rl	30	24	6
Bs + Pf	24	9	15
Bs + Rl	27	15	12
Pf + Rl	19	9	10
Bs + Pf + Rl	29	13	16
Total	233	135	98

Table 2. The number of seeds treated and the numbers of seeds that did and did not germinate for each treatment group are depicted in this table. Seed numbers of the entire experiment for each of these categories are totaled at the bottom of this table.

Data Analysis

The pictures were analyzed using the imaging software ImageJ. We obtained digital measurements with this program by first using the 1.3 cm grid lines on the plates to set a scale of approximately 1950 pixels per cm and then using a freehand line tool to trace the section of the root that needed to be measured. We added the lengths of each root segment together to calculate the total root length for each seedling. The number of root branches, characterized by how many times the root diverged, was also recorded. We included the length of each root branch in the total length measurements. We then analyzed these data values to determine if the differences in total length or the number of root branches were statistically significant among the various treatments. Using a one-way Analysis of Variance (ANOVA) test, we compared the means of each group collectively. With a significance level of 0.05, we identified the *p*-values, which indicated whether the different treatment groups had an effect on plant growth. *P*-values less than 0.05 denoted statistical

significance between the differences in growth among the different groups.

ACKNOWLEDGEMENTS

We would like to thank Richard Kurtz, Janine Cupo, Davide Bulgarelli, Susanna Harris, and Harvard's Journal of Emerging Investigators Editors for their guidance and support throughout this research process.

Received: June 11, 2020
Acccepted: September 29, 2020
Published: October 25, 2020

REFERENCES

1. Hincks, Joseph. "The World Is Headed for a Food Security Crisis. Here's How We Can Avert It." *Time*, 28 Mar. 2018, www.time.com/5216532/global-food-security-richard-deverell/.
2. Schiffman, Richard. "Why It's Time to Stop Punishing Our Soils with Fertilizers." *Yale Environment 360*, 3 May 2017, www.e360.yale.edu/features/why-its-time-to-stop-punishing-our-soils-with-fertilizers-and-chemicals.
3. Arsenault, Chris. "Only 60 Years of Farming Left If Soil Degradation Continues." *Scientific American*, 5 Dec. 2014, www.scientificamerican.com/article/only-60-years-of-farming-left-if-soil-degradation-continues/.
4. Bhattacharyya, P. N., & Jha, D. K. "Plant Growth-Promoting Rhizobacteria (PGPR): Emergence in Agriculture." *World Journal of Microbiology and Biotechnology,* vol. 28, 2012, pp. 1327-1350.
5. Woo, Sheridan L., & Pepe, Olimpia. "Microbial Consortia: Promising Probiotics as Plant Biostimulants for Sustainable Agriculture." *Frontiers in Plant Science*, vol. 9, 2018, pp. 1801.
6. Saeed, Kamil, et al. "Effect of Bio-Fertilizer and Chemical Fertilizer on Growth and Yield in Cucumber (Cucumis sativus) in Green House Condition." *Pakistan Journal of Biological Sciences,* vol. 18, no. 3, 2015, pp. 129-134.
7. Hu, Jie, et al. "Probiotic *Pseudomonas* Communities Enhance Plant Growth and Nutrient Assimilation via Diversity-Mediated Ecosystem Functioning." *Soil Biology & Biochemistry*, vol. 113, 2017, pp. 122-129.
8. Koornneef, Maarten., and Meinke, David. "The Development of *Arabidopsis* as a Model Plant." *The Plant Journal*, vol. 61, 2010, pp. 909-921.
9. Finch-Savage, William E., and Bassel, George W. "Seed Vigour and Crop Establishment: Extending Performance Beyond Adaptation." *Journal of Experimental Botany*, vol. 67, no. 3, 2016, pp. 567-591.
10. Adesemoye, Tony O., et al. "Comparison of Plant Growth-Promotion with *Pseudomonas aeruginosa* and *Bacillus subtilis* in Three Vegetables." *Brazilian Journal of Microbiology*, vol. 39, 2008, pp. 423-426.
11. Yanni, Youssef G., and Dazzo, Frank B. "Enhancement of Rice Production using Endophytic Strains of *Rhizobium leguminosarum* bv. Trifolii in Extensive Field Inoculation Trials within the Egypt Nile Delta." *Plant and Soil,* vol. 336, 2010, pp. 129-142.
12. Gopalakirshnan, Subramaniam, et al. (2015). "Plant Growth Promoting Rhizobia: Challenges and Opportunities." *3 Biotech*, vol. 5 no. 4, 2015, pp. 355-357.
13. Niu, Ben, et al. "Simplified and Representative Bacterial Community of Maize Roots." *Proceedings of the National Academy of Sciences of the United States of America,* vol. 114, no. 12, 2017.
14. Nathoo, Naeem, et al. "A Hydroponic Co-Cultivation System for Simultaneous and Systematic Analysis of Plant/Microbe Molecular Interactions and Signaling." *Journal of Visualized Experiments*, vol. 125, 2017.
15. Bochicchio, Rocco, et al. "Root Architecture and Morphometric Analysis of *Arabidopsis thaliana* Grown in Cd/Cu/Zn-Gradient Agar Dishes: A New Screening Technique for Studying Plant Response to Metals." *Plant Physiology and Biochemistry*, vol. 91, 2015, pp. 20-27.

JOURNAL OF
EMERGING INVESTIGATORS

Article

Developing a Method to Remove Inorganic Arsenic from Rice with Natural Substances

Kayla K. Mukai[1] & Yvonne L. Chan[1]
[1] 'Iolani School, Honolulu, Hawaii.

SUMMARY

Many developing nations struggle with arsenic (As) exposure due to the over-pumping of groundwater wells, which results in higher concentrations of As in agricultural systems. In Bangladesh, there has been an ongoing crisis of As-related health issues, such as heart disease and various cancers, primarily due to the consumption of As-contaminated water and rice. Prior research has mainly focused on removing As from water, but there has been little research on removing As from rice, even though rice contains dangerous amounts of As and Bangladeshis have a heavily rice-based diet. We focus on developing a method to remove inorganic As from rice using natural substances. We developed a method utilizing activated charcoal and bentonite clay to lower levels of As in rice using these natural substances. We adapted a procedure to use the Quick Sensafe Rapid Arsenic Test Kit to measure As before and after treatment while simulating cooking conditions. Activated charcoal was the most effective, as it was able to remove 75% of As when used alone. Therefore, this method could be utilized by countries like Bangladesh to lower As intake and incidences of As-related health issues.

INTRODUCTION

Bangladesh is facing a major public problem of arsenic (As) contamination due to their As-polluted tube wells and groundwater systems (1). Unfortunately, this As contamination has spread to their food resources, as As has accumulated at significant concentrations in locally grown rice (2). Since rice is grown under flooded conditions, toxic inorganic As from polluted groundwater builds up in large amounts around the root and is absorbed in high concentrations into the grain which is then consumed (2). Rice plays a key role in Bangladesh As exposure (3). Bangladeshis have a heavily rice-based diet, and many individuals are inadvertently consuming high levels of As. In West Bengal and Bangladesh, rice is the main source of As exposure after water and on average, Bangladeshis consume 450 to 600 grams of rice each day (4). Consumption of inorganic As is a cause of numerous health issues, including skin lesions, heart disease, and skin, bladder, and lung cancers (5). Therefore, we looked at the As in rice that people consume and developed and tested methods for removing As from rice.

In a study conducted by the Health Effects of Arsenic Longitudinal Study (HEALS), participants from Bangladesh were tracked for over a decade, and those exposed to the highest concentrations of As displayed an increase of nearly 70% in all-cause mortality (6). HEALS also reported that even low-level exposures to As increased mortality rates, showing that any amount of As intake could potentially be deadly. There have also been recent studies which correlate high As exposure with higher levels of illnesses and death. In Matlab, Bangladesh participants were monitored for thirteen years, and the study concluded that higher consumption of As increases mortality risks among young adults and that As poisoning is prevalent in Bangladeshi society today (7).

Current methods to lower inorganic As concentration involve washing the rice multiple times, discarding the contaminated water, and then boiling the rice with extra water to lessen the reabsorption of As back into the grain (8). However, this method is not completely effective. One study showed that only up to 57% of total As levels could be removed with low-As water washing and boiling with extra water, and 28% could be removed with low-As water washing alone (9). Consumption of the remaining inorganic As is potentially dangerous (10). Thus, complete As removal from rice is an ongoing issue.

Inorganic As levels in rice vary based on the type of rice and where it is grown. Brown rice absorbs more inorganic As because As accumulates in the aleurone layer of the rice grain, which gives brown rice its color (10). Also, As concentration increases significantly as the size of the rice grain increases. Therefore, large-grain rice is more contaminated. That is why we used large-grain brown rice from India that contained significant levels of inorganic As.

The goal of this study was to develop a method to remove as much inorganic As (arsenite, AsIII and arsenate, AsV) from rice as possible using the natural substances activated charcoal and bentonite clay. Activated charcoal and bentonite clay have similar properties, including their ability to adsorb heavy metals due to their negatively-charged, porous surfaces. Particularly, activated charcoal has a large surface area that allows for the physical adsorption of dissolved substances, such as inorganic As, from liquids (11). The use of natural substances like activated charcoal and bentonite clay to lower inorganic As levels from solution is effective, but their abilities to lower inorganic As in solids, such as rice, have not been tested. We hypothesized that inorganic As levels in rice would be lowered after the addition of these

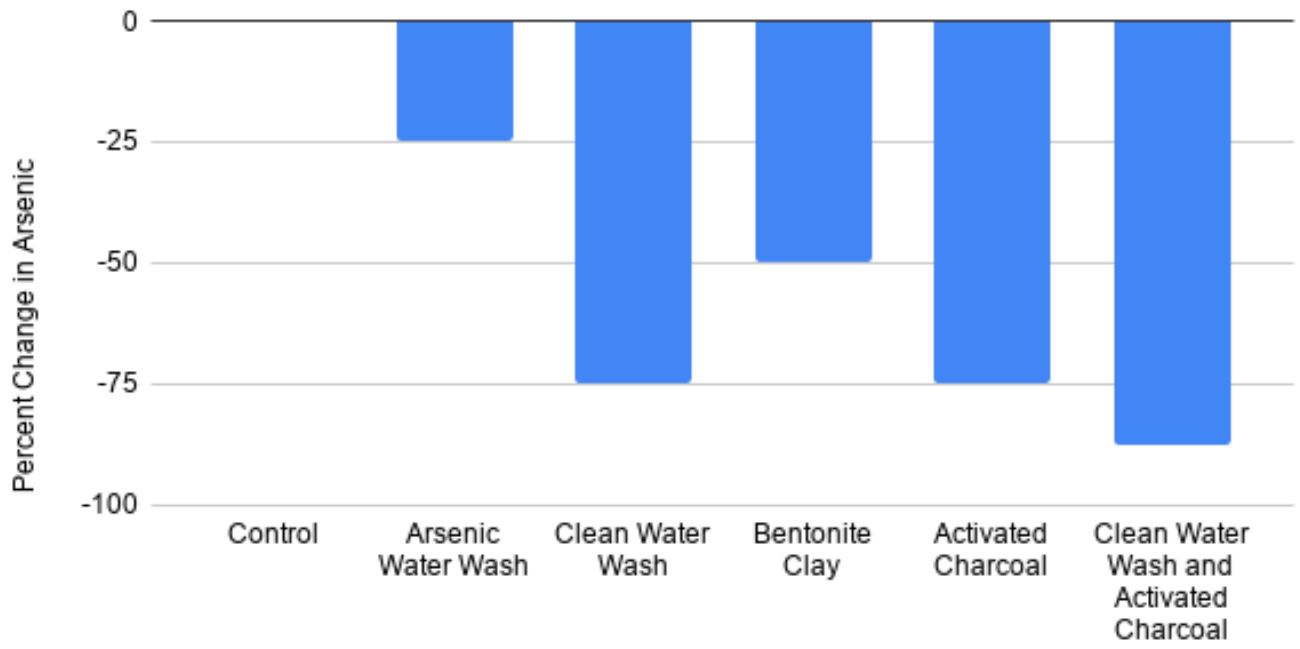

Figure 1: Percent decrease in As in rice after various removal methods. We tested As water wash, clean water wash, bentonite clay, activated charcoal, and combined clean water wash and activated charcoal. Results are reported as percent decrease compared to control (no wash). Each discrete color result of the test kit indicates an As concentration within a 10 ppb range, which was consistent among each of the three trials for each method; thus, there are no error bars.

natural substances. We showed that activated charcoal and bentonite clay can remove As from rice and potentially help lower levels of As intake. Our research could benefit those living in Bangladesh, as it could help remove inorganic As from their diets and lower occurrences of As-related health issues.

RESULTS

The effectiveness of water wash methods, bentonite clay, and activated charcoal were examined for their efficiency in removing As from rice. For all tests, the rice was cooked for twenty minutes, dried in an incubator until no moisture remained, and homogenized with a coffee grinder. The powdered rice was then added to water and heated to extract the As. The resulting substance was then cooled and As levels were measured using the Sensafe Quick Arsenic field test kit. This test kit utilizes a patented, modified Gutzeit method for the reaction, which creates arsine gas and displays results on a mercuric bromide strip. These kits are reliable, fast, inexpensive, and produce results comparable to higher end machines (12). The Sensafe kit provided an Easy Read color chart that is used to match the color formed on the mercuric bromide strip with the amount of inorganic As in the initial sample. This As test kit was created for water, soil, and wood analysis, however, which is why we adapted this procedure to have the ability to measure As in rice.

We used two water wash methods: As water wash and clean water wash. The As water wash method, which simulates Bangladesh conditions by using a 10 ppb As solution, lowered the original 40 ppb of inorganic As to 30 ppb. The clean water wash method, which used As-free water, lowered the 40 ppb to 10 ppb of inorganic As (**Figure 1**).

The usage of bentonite clay to remove inorganic As from rice is effective but requires a method to prevent the bentonite clay powder from mixing with the rice. The usage of a coffee filter with 10 grams of bentonite clay powder lowered the concentration of As from 40 ppb to 20 ppb (**Figure 1**).

The usage of a five-gram activated charcoal brick lowered the original 40 ppb to 10 ppb (**Figure 1**). In addition to using activated charcoal on its own, we combined this method with the clean water wash method to attempt to remove even more inorganic As. The results of this were favorable, as the two methods combined are able to remove the most inorganic As, lowering the concentration of As to 5 ppb (**Figure 1**).

DISCUSSION

Our results support the hypothesis that natural substances, activated charcoal and bentonite clay, are able to remove inorganic As from rice. The various methods differ in the removal of As from the rice. Results appeared consistent throughout all three trials because the kit used for these experiments is not very sensitive (**Figure 1**). The combination of the clean water wash and activated charcoal method removed the most As, followed by the activated charcoal method tied with the clean water wash method, then the bentonite clay method, and finally the As water wash method which removed the least.

By using the As water wash method, 25% of inorganic As can be removed (**Figure 1**). Previous research found that washing rice with low-As water can remove up to 28% of

inorganic As (9). This supports the validity of our measurement procedure because the results closely align.

The method that resulted in the second-lowest level of inorganic As removal of 50% is bentonite clay (**Figure 1**). Bentonite clay is not the most effective at removing inorganic As, as only half could be removed when using the coffee filter method. This could be due to the usage of the coffee filter pouch that could have prevented exposure of inorganic As to the bentonite clay's surface. Unfortunately, this coffee filter method is necessary to prevent the combination of the clay powder with the rice. Zahra *et al*. reports that bentonite clay is able to remove up to 95% of inorganic As from wastewater (13), but our study's results show that only 50% of inorganic As can be removed from rice (**Figure 1**). However, this result of 50% removal still supports our hypothesis that bentonite clay can remove inorganic As.

Two methods resulted in removal of 75% of the inorganic As: the clean water wash and activated charcoal treatment (**Figure 1**). The clean water wash method is effective and results in a removal of 75% of inorganic As. However, this method can only be utilized by those with access to clean water, a potential limitation to those in Bangladesh with contaminated tube wells and groundwater systems (1). Activated charcoal is also effective and results in a removal of 75% of inorganic As. These results can be supported with previous research that reports that activated charcoal is able to adsorb AsV by 2.5% of its weight and AsIII by 1.2% of its weight from solution (14).

The clean water wash method combined with the activated charcoal method works the best to remove the most inorganic As. The combination of these methods is tested because both show the most potential for high As removal. This combination results in a removal of 87.5% of inorganic As, which is the greatest As removal percentage observed in our study (**Figure 1**). For those with access to clean water, this method would work. Future studies should investigate whether the combination of the As water wash method with the activated charcoal method is effective for people in Bangladesh. However, the activated charcoal method alone is still effective, as it is able to remove 75% of inorganic As without the clean water wash. Also, activated charcoal is very convenient due to its ability to remain activated for 1-2 months and its simplicity to use. In conclusion, activated charcoal would be the most applicable and effective for Bangladeshis.

One possible limitation is that the maximum amount of inorganic As that could be extracted from the plain Indian brown rice using the kit was 40 ppb. This is most likely not all of the As found in the rice. However, by keeping the procedure the same for testing the other methods, this gave us a baseline value to which the experimental results could be compared. Additionally, the test strips provided in the kit only had colors in 10 ppb increments to measure As levels, so there could have been possible errors in the precision of the measurements, as we were unable to measure removal to the nearest percent. To limit error, we had at least three trials per method and ensured that the colors on the test strip appeared consistent.

The findings resulting from our study could potentially help those in Bangladesh lower their As intake, therefore possibly lowering the occurrences of As-related health issues and death (6). In the future, we should test more methods, such as boiling with extra water, and test more types of rice, obtaining rice from Bangladesh or testing rice from the United States. Also, we should vary the amounts of adsorbent substance used in each experiment in order to determine if that could possibly have an effect on removal capabilities. Additionally, we could look at more ways to make these methods more accessible to those in Bangladesh by possibly creating activated charcoal out of easily obtainable resources such as coconut shells. As of now, the results are favorable because the initial goal to develop a method to remove inorganic As from rice using natural substances is met.

MATERIALS AND METHODS

Measurement of Arsenic Levels

We used the Sensafe Quick Arsenic field test kit (Industrial Test Systems Part No. 481396) to measure As levels. This test kit utilizes a patented reaction, a modified Gutzeit method, which creates arsine gas and displays results on a mercuric bromide strip. In this method, the inorganic As compounds in the sample are converted to arsine (AsH_3) gas, which is a product of the reaction of zinc and tartaric acid. Ferrous and nickel salts are added to speed up this reaction. The arsine gas reacts with the mercuric bromide on the test strip to form mixed mercury halogens like AsH_2HgBr to create a yellow or brown color. The test kit provided an Easy Read color chart with low sensitivity that was used to match the color with the amount of inorganic As in the initial sample. This As test kit was created for water, soil, and wood analysis, however, which is why we adapted this procedure to have the ability to measure As in rice.

Preparation & Cooking

To simulate cooking conditions, 25 grams of large-grain brown rice from India was cooked in 60 mL of water for 20 minutes and then dried in an incubator at 28ºC to return it back into its dried form that could be measured for As with the Sensafe Quick Arsenic field test kit. Since the field test kit that we use detects levels of inorganic As in water, the rice had to be homogenized. First, a coffee grinder was used to grind the rice into a powder, and then 10 grams of this rice powder was placed into 100 ml of water. This mixture was heated on a consistent medium-high setting for five minutes to further extract As. After cooling, the mixture was ready to be tested using the test kit.

First, using the aforementioned procedure, we took multiple baseline samples of the As concentration in the rice alone as a control. After finalizing the control measurement procedure, we continued on testing the other methods. The current water wash method was tested by swirling the rice in

water for ten seconds and discarding and changing the water three times before cooking. Water washing was conducted before the rice was cooked, and then the rest of the procedure was followed. In an attempt to develop a new method to remove As that utilized activated charcoal or bentonite clay, we cooked the rice with these natural substances by either placing it in alone or putting it inside of a permeable coffee filter pouch. Specifically, we utilized five grams of activated charcoal and ten grams of bentonite clay powder within a permeable pouch throughout the entire cooking duration. Activated charcoal was in a solid brick form and placed in direct contact with the rice. The bentonite clay was in powdered form and placed inside of a coffee filter pouch. For all tests, the rice was cooked for twenty minutes, dried in an incubator until no moisture remained, and homogenized with a coffee grinder. Three trials of each method were conducted, and results appeared consistent.

Data Analysis

To analyze the data, we compared the results of the experimental groups to the results of the control group. **Figure 1** was prepared by finding percent decrease compared to the control group. Each discrete color result of the test kit indicates an As concentration within a 10 ppb range.

ACKNOWLEDGEMENTS

The authors are grateful for the assistance of Kelly Mukai for help in the lab and ʻIolani School for use of the Wet Lab and funding the purchase of the As test kit.

Received: June 18, 2020
Accepted: October 16, 2020
Published: October 26, 2020

REFERENCES

1. Ahmad, Sk Akhtar *et al*. “Arsenic contamination in groundwater in Bangladesh: implications and challenges for healthcare policy.” *Risk management and healthcare policy,* vol. 11 251-261. 30 Nov. 2018, doi:10.2147/RMHP.S153188
2. Zavala, Yamily J., and John M. Duxbury. “Arsenic in Rice: I. Estimating Normal Levels of Total Arsenic in Rice Grain.” *Environmental Science & Technology*, vol. 42, no. 10, 2008, pp. 3856–3860., doi:10.1021/es702747y.
3. Melkonian, Stephanie, *et al*. "Urinary and dietary analysis of 18,470 Bangladeshis reveal a correlation of rice consumption with As exposure and toxicity." *PloS one*, 8.11 (2013).
4. Shakoor, Muhammad Bilal, *et al*. "Recent Advances in As Accumulation in Rice." *Advances in Rice Research for Abiotic Stress Tolerance*. Woodhead Publishing, 2019. 385-398.
5. FDA, US, and US Food and Drug Administration. "As in rice and rice products risk assessment report." (2016).
6. Argos, Maria, *et al*. "As exposure from drinking water, and all-cause and chronic-disease mortalities in Bangladesh (HEALS): a prospective cohort study." *The Lancet*, vol. 376, no. 9737, 2010, pp. 252-258.
7. Rahman, Mahfuzar, *et al*. "As exposure and young adult's mortality risk: A 13-year follow-up study in Matlab, Bangladesh." *Environment international,* 123 (2019): 358-367.
8. Raab, Andrea, *et al*. "Cooking rice in a high water to rice ratio reduces inorganic As content." *Journal of Environmental Monitoring*, 11.1 (2009): 41-44.
9. 9. Sengupta, M. K., *et al*. "As burden of cooked rice: traditional and modern methods." *Food and Chemical Toxicology,* 44.11 (2006): 1823-1829.
10. Schmidt, Charles W. "Low-dose As: in search of a risk threshold." (2014): A130-A134.
11. Mohammad-Khah, A., and R. Ansari. "Activated charcoal: preparation, characterization and applications: a review article." *Int J Chem Tech Res,* 1.4 (2009): 859-864.
12. Bralatei, Edi, *et al*. "Detection of inorganic As in rice using a field test kit: a screening method." *Analytical chemistry,* 87.22 (2015): 11271-11276.
13. Zahra, Naseem, *et al*. "Removal of As from wastewater using bentonite." *Bangladesh Journal of Scientific and Industrial Research* 44.1 (2009): 81-86.
14. Eguez, Hugo Ernesto, and Eung Ha Cho. "Adsorption of As on activated charcoal." *Jom* 39.7 (1987): 38-41.

Sponsorship

Editor's Circle $10,000+

ALLOY THERAPEUTICS

Patron $5,000+

NEW ENGLAND BioLabs® Inc.
enabling technologies in the life sciences

PORTFOLIOS WITH PURPOSE®

Institutional Supporters

HARVARD UNIVERSITY

HARVARD MEDICAL SCHOOL

Tufts UNIVERSITY

Charitable Contributions

We need your help to provide mentorship to young scientists everywhere.

JEI is supported by an entirely volunteer staff, and over 90% of our funds go towards providing educational experiences for students. Our costs include manuscript management fees, web hosting, creation of STEM education resources for teachers, and local outreach programs at our affiliate universities. We provide these services to students and teachers entirely free of any cost, and rely on generous benefactors to support our programs.

A donation of $30 will sponsor one student's scientific mentorship, peer review and publication, a six month scientific experience that in one student's words, 're-energized my curiosity towards science', and 'gave me confidence that I could take an idea I had and turn it into something that I could put out into the world'. **If you would like to donate to JEI, please visit https://emerginginvestigators.org/support, or contact us at questions@emerginginvestigators.org.** Thank you for supporting the next generation of scientists!

'Journal of Emerging Investigators, Inc. is a Section 501(c)(3) public charity organization (EIN: 45-2206379). Your donation to JEI is tax-deductible. '

emerginginvestigators.org