Using Linear Regression Correlational Analysis to Understand the Relationship Between Cyanobacterial Harmful Algal Blooms and Parkinson's Disease Prevalence in Various Regions of the US

Esha Agarwal

Abstract— **Parkinson's Disease (PD) is a movement/behavioral disorder that affects millions worldwide. Due to its neurodegenerative nature, PD's induction and prognosis remain unknown; thus, it has an uncharacteristically high misdiagnosis rate. Scientists are investigating genetic predispositions and environmental factors that may result in PD. BMAA, a neurotoxic byproduct of Cyanobacterial Harmful Algal Blooms (CyanoHABs), has been found to cause neurodegeneration [1]. To better understand this relation, this study aims to understand whether there is a correlation between CyanoHABs and PD prevalence in the US, especially in a period of climate change. After finding CyanoHAB prevalence data of four states that fit given criteria from the past 4-5 years, PD prevalence was obtained by determining a PD state proportion and estimating the PD prevalence in each state. A Python algorithm was designed to conduct a linear regression correlational analysis on the data, and the correlation and determination coefficients were determined. Based on the coefficients of determination, each state was found to have a moderate-high correlation between the independent variable of CyanoHABs and dependent variable of PD and a positive correlation between CyanoHABs and PD was supported. Limitations for this study included PD prevalence estimation and lack of CyanoHAB data, both of which limited the duration and accuracy of this study. In the future, using satellite imagery and contacting medical centers directly can overcome these issues. Given that the elderly population is rising in many developed nations and increasing climate change and CyanoHABs in the future, it is becoming extremely important to better understand the disease [2],[3].**

Index Terms— **Cyanobacteria/Cyanobacterial Harmful Algal Blooms, Linear Regression Correlation Analysis, Parkinson's/Parkinsonian Disease, Remote Sensing & Satellite Imaging.**

I. INTRODUCTION

Parkinson's Disease is a multisystem movement disorder characterized by the three cardinal symptoms of bradykinesia, or slowed movement; tremor, or the shaking of hands and feet; and muscle rigidity, or body stiffness. Because much information about PD remains unknown, PD has an unusually high misdiagnosis rate, reaching almost 26% [4]. This study has the ability to give scientists further insight into the disease because this study is one of the only papers to investigate how PD may be linked to Cyanobacterial Harmful Algal Blooms

Esha Agarwal, Enloe High School, US

and climate change, imminent issues in today's society. In addition, the population of various developed countries is aging quickly: "Over the next fifteen years, the proportion of the population ages 65 and older and ages 85 and older will rise dramatically...in 2030, the proportion of elderly people in the U.S. population will only reach levels already evident today in Germany, Italy, and Japan" [5, p. 6]. Given that the risk for PD and neurodegenerative disease increases greatly with age, this study has great potential to contribute to the future of the disease.

CyanoHABs and BMAA. Harmful Algal Blooms, or HABs, are a water-based phenomena that can occur in all marine environments. Essentially, HABs are occurrences in which there is an overproduction of algae and similar organisms that results in harmful effects on the health of the marine environment, ecosystem, and human activities. Algal blooms are a normal seasonal occurrence; however, due to increasing global warming and climate change, blooms are becoming larger and more frequent [6]. CyanoHABs, or Cyanobacterial Harmful Algal Blooms are similar to HABs, and as described in "Measurement of Cyanobacterial Bloom Magnitude using Satellite Remote Sensing," by Mishra et al. [7], "[CyanoHABs] are a serious environmental, water quality and public health issue worldwide because of their ability to form dense biomass and scum and to produce toxins such as neurotoxins" [7, p. 1]. CyanoHABs are composed primarily of cyanobacteria, neurotoxin-photosynthetic bacteria that survive in marine environments [1]. Cox et al. [1] concluded that all known forms of cyanobacteria, including those that live in varied types of water bodies, have the capacity to produce BMAA, or beta-Methylamino-L-alanine, a neurotoxin that has been linked to causing Parkinson's Dementia Complex/Amyotrophic Lateral Sclerosis (PDC/ALS) and similar neurodegenerative diseases. BMAA was initially not considered to be associated with neurodegenerative disease. Recently, the link between BMAA and neurodegenerative disease was determined when large numbers of the Chamorro people of Guam developed PDC/ALS. It was found that their diet consisted of cycad seeds and cycad seed products, as well as flying foxes; each of these food products contained high levels of BMAA, and through biomagnification, the BMAA concentration in the Chamorro people reached dangerously high levels, resulting in neurodegeneration [1].

The overproduction of algae, dinoflagellates, and

cyanobacteria that occur in HABs and CyanoHABs result from eutrophication, a form of nutrient pollution in which the excess nutrients of nitrogen and phosphorus enter and stimulate the photosynthetic organism growth in a body of water. One study, by Paerl and Whitall [8], environmental scientists, found that one of the largest sources of nitrogen that enters water bodies and results in eutrophication and HABs is AD-N, or Atmospheric Deposition Nitrogen. AD-N includes "all biologically available forms of dissolved inorganic nitrogen (NOx NO-', NO-', emitted from fossil fuel combustion and biomass burning; NH3/NH', a volatilization product of agricultural waste, fertilizers, decomposition, biomass burning) and dissolved organic nitrogen" [8, p. 307]. HABs and CyanoHABs are harmful in a variety of ways. Because the nutrient pollution results in an overproduction of algae and photosynthetic organisms, this layer of "scum" can produce harmful neurotoxins such as BMAA and other phycotoxins, deplete the water body's dissolved oxygen content, and prevent penetration of sunlight into the water body, disrupting the ecosystem [9]. Marine animals and fish can experience gill damage and even death from the toxins, and marine birds may experience hypothermia from the surfactant-like proteins created by the algae that wear away at the insulation and waterproofing provided by feathers [9]. Humans are also affected by HABs and CyanoHABs, industries such as fishing and tourism are negatively impacted by the disrupted ecosystems, and the neurotoxins in the water can affect human health if they are ingested or absorbed through skin [9].

A. Establishing the Existing Correlation Between BMAA and PD

Study-Analysis 1. Parkinson's Disease has long been thought to be correlated with environmental factors and toxins [10]. As mentioned above, one of the earliest environmental correlations was between PDC and BMAA; it was discovered amongst the Chamorro people in Guam in the early 1950s. Since then, BMAA has been thoroughly investigated and a strong correlation between Parkinsonian disease and BMAA has been found [11]. To determine and further investigate the correlation, as well as how BMAA functions in the brain, Arif and his fellow researchers [11] treated Wistar neonatal rats with 1 mM of BMAA for 2 hours. Neonatal rats were used because the toxin can pass younger rats' blood-brain barrier more easily. The researchers found that BMAA did increase phosphorylation in most amino acids that produce the tau protein. This is significant as increased phosphorylation leads to the formation of tau oligomers, which leads to Lewy Body formation (protein-aggregates) in and around the nervous system and brain, resulting in PD-like symptoms. Researchers also studied how PP2A, a regulatory dephosphorylatory protein, interacted with the hyperphosphorylation of tau after BMAA [11]. It was found that BMAA interacts directly with PP2A, most likely inhibiting PP2A's signaling pathways, preventing PP2A from dephosphorylating the now hyperphosphorylated tau protein, and preventing homeostasis. These findings are significant and effectively show a correlation between BMAA and the presence of PD biomarkers such as Lewy bodies and changes in tau pathology as they lead to Parkinsonian diseases.

Study-Analysis 2. The study conducted by Arif et al. [11] is not the only study that finds a correlation between BMAA and the development of Parkinsonian diseases. A study conducted by Cox [10] uses Vervet monkeys to explain how BMAA enters the body systems, affects different neural pathways, and leads to neurodegenerative disease, "BMAA rapidly crosses the blood-brain barrier, where it is captured by the central nervous system…mistaken by cellular machinery for L-serine and be misincorporated into proteins, leading to protein misfolding, aggregation and subsequent apoptosis" [10, p. 2]. The study also discusses how BMAA affects different neurological pathways: "our data suggest that chronic dietary exposure to BMAA triggers both the NFT and ß-amyloid pathways" [10, p. 2]. This describes how BMAA triggers the NFT and ß-amyloid pathways, and leads to the buildup of plaque, which is a similar substance to Lewy Bodies, a biomarker commonly found in Parkinson's patients. In conclusion, the study finds that extended exposure to BMAA through consumption does in fact lead to ALS/PDC because the experiment satisfies Koch's three Postulates. Once again, this study is important because it effectively demonstrates the relationship between BMAA and Parkinsonian diseases using the Vervet monkey animal model.

B. Investigating the Gap

The initial question I had was to determine whether BMAA was correlated with PD or Parkinsonian diseases. Upon getting acquainted with the discipline-specific literature, I realized that this gap had already been filled. It is important to note that although BMAA and Parkinsonian diseases have been correlated, the reservoir of BMAA, Cyanobacterial Harmful Algal Blooms (CyanoHABs), have not been directly correlated to the Parkinsonian diseases. To do so would be a non-sequitur, since drawing a direct correlation would be incorrect [12]. After understanding that the initially-identified gap had been filled, I began investigation in other directions, while keeping in mind the topic of Parkinson's Disease and my research goal of helping scientists and patients with PD or other neurological diseases better understand the pathogenesis of PD. I learned of a relative of cyanobacteria, melainabacteria, that lives in the gut microbiome of the human body but lacks the internal machinery of cyanobacteria, rendering it unable to produce a toxin such as BMAA [13]. While melainabacteria does evoke many interesting questions about the role of the gut in the prognosis and development of PD, due to the lack of peer-reviewed sources--according to Nunes-Costa et al. [13], the lack of the internal machinery that produces a toxin caused many scientists to disregard it--the gap for studying melainabacteria was too large, and therefore difficult to pursue.

After extensive research into the relation of PD and environmental toxins, I realized that while many have studied and determined the relationship between BMAA and PD, articles that determined a correlation between the source of

BMAA, CyanoHABs, and PD had not been encountered. Thus the research question for this study is: How have Cyanobacterial Harmful Algal Blooms affected the prevalence rates of Parkinson's Disease across various regions of the US in the last 4-5 years?

C. Significance and Justification of this Study

While researchers are aware of the pathogenesis of PD, they remain unsure of the origins, and are therefore investigating environmental factors such as neurotoxins like BMAA and pesticides such as Paraquat and Rotenone, as well as genetics and genes such as SNCA, PINK1, LRRK2 and more. This study aims to help scientists and researchers, as well as those suffering from neurological diseases like PD, gain further insight into the disorder. Climate change is known to be on the rise, "atmospheric concentrations of greenhouse gases are 40 percent higher [...] 400 parts per million today, compared with 280 parts per million in the 1780s [...] three-quarters of that rise occurring over the past 45 years" [14, p. 69]. In addition, climate change's acceleration results in the increase of the creation of HABs. Climate change is resulting in higher temperatures and many algae and marine photosynthetic organisms flourish in higher temperatures. In addition, global warming has lead to increased carbon dioxide, which is once again helpful to the growth of algae and the creation of HABs, "Climate change directly modifies temperature, which in turn affects HAB growth rates [...] pigment content […and] photosynthesis...many cyanobacterial and dinoflagellate species, including HAB species, prefer warmer temperature conditions [...] The toxicity of many harmful algal species also increases with warming" [15, p. 64].

D. Summary

In summary, Parkinson's Disease is a movement and behavioral multisystem disorder that currently affects millions across the globe, and due to increasing climate change and aging populations in many developed countries, has the potential to become a rampant and widespread disorder. Environmental factors have long been suspected for the induction of PD, and BMAA, a neurotoxin produced in Cyanobacterial Harmful Algal blooms--these are phenomena that have been occurring due to increased pollution and climate change--has been found to be linked to both PDC and PD. This study aims to better understand the relation between the neurodegenerative disease PD and CyanoHABs.

II. METHODS

To further investigate how climate change and PD prevalence are correlated through the increasing occurrence of CyanoHABs, this study utilizes a double-step quantitative objective retrospective approach, procuring data from publicly available datasets to determine the prevalence of both CyanoHABs and PD in various states in the past few years and then performing a linear regression analysis to determine the potential presence of a correlation. According to Dr. Dean R Hess [16], a retrospective study, also known as a retrospective cohort study, "uses existing data that have been recorded for reasons other than research" [16, p. 1171]. This study has two steps of quantitative data collection, determining the prevalence rates of CyanoHABs and then of

PD. A correlational analysis similar to that performed by Isenstein et al. [17] in "Multispectral Remote Sensing of Harmful Algal Blooms in Lake Champlain, USA" will also be performed. In the aforementioned paper, the use of satellite imagery to better understand the properties of the water body is described:

Satellite remote sensing is especially useful for bloom monitoring...can detect...photosynthetic pigments such as chlorophyll a (chl-a)...satellite data sets presents many opportunities for retrospective analysis...The spatial and temporal distribution derived from satellite imagery…can be used to implement effective strategies [17 p. 2271].

Because satellite imagery and remote sensing have been used by multiple peer-reviewed studies and have the ability to effectively characterize a great number of water bodies, this justified method was used in this study to determine the prevalence or annual temporal frequency of CyanoHABs. Because satellite images give many details about a water body, such as chlorophyll concentration and sea surface temperature, initially, the chlorophyll concentration was utilized to characterize the bodies of water. However, looking at hundreds of individual satellite images proved to be time-consuming and difficult to complete with limited available software. Therefore, determining if there is a correlation between chlorophyll content of various bodies of water and PD rates remains a future direction. Instead, the prevalence of CyanoHABs in various states was determined from state government websites and their satellite maps that have tracked the growing issue of CyanoHABs in the past 4-5 years. The study done by Isenstein et al. [17] also conducted a linear regression correlational analysis in R; this study utilizes the Python scipy, numpy, matplotlib.pyplot, and math libraries to perform the linear regression analysis to determine if there is a correlation between CyanoHABs and PD. The PD prevalence rates for various states were gathered from past retrospective studies. This method, in Figure 1, involves the utilization of a double-step quantitative retrospective methodology as it includes the collection of CyanoHAB and PD prevalence rates from government websites, and a correlational linear regression statistical analysis done to determine the presence of a correlation.

Figure 1
Proposed Methodology
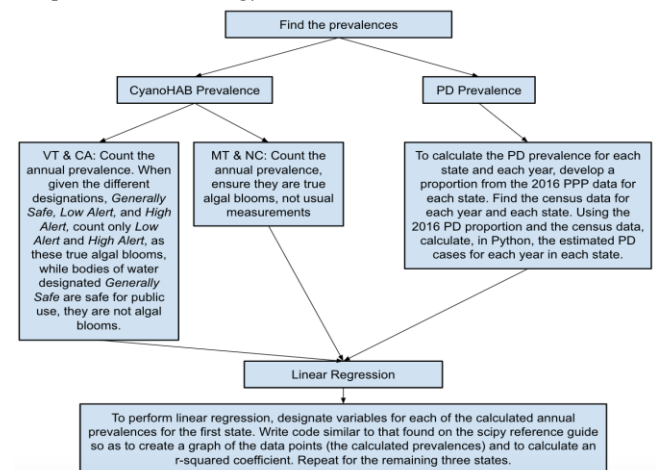
A. Acquiring CyanoHAB Data

In order to determine CyanoHAB prevalence in various states, the government website entitled "State HABs Resources" and each of the embedded links were studied in great depth. Investigation continued until states of the correct criteria were met. The criteria was as follows: half of the states must be coastal while the other half of the states must be landlocked, the states must have CyanoHAB prevalence data from between the past 4-5 years available, and there must be at least 100 identified PD cases in each state. The states of North Carolina, California, Vermont, and Montana all met this criteria; California and North Carolina are coastal states while Vermont and Montana are land-locked, all of the states have available prevalence data from 2017 and 2018, and each state has over 100 confirmed PD cases.

Collecting the Data. To determine the annual CyanoHAB prevalence in North Carolina, Montana, Vermont, and California from 2014-2019, 2017-2020, 2014-2018, and 2016-2020 respectively, the CyanoHAB total for each year was found. The exact government sites used are listed in the Index.

Special Cases. Each state has a slightly different method of recording the CyanoHAB prevalence. Vermont monitors its water bodies closely, collecting samples from each water body weekly. Each sample is classified as either *Generally Safe, Low Alert,* or *High Alert*, depending on the perilousness of the water [18]. In this study, only water bodies marked as *Low Alert* or *High Alert* were added to Vermont's annual CyanoHAB prevalence. California too marks its water bodies as either *Generally Safe, Low Alert,* or *High Alert*; only water bodies marked as *Low Alert* or *High Alert* were added to California's annual CyanoHAB prevalence [19]. In Montana and North Carolina, the annual prevalence of CyanoHABs in the state was counted simply based on reports; cyanobacteria can live in all forms of water bodies and all cyanobacteria produce neurotoxins; thus, all algal blooms were included [1].

B. Acquiring PD Prevalence Data

While it was challenging to find CyanoHAB prevalence data, it proved much more difficult finding recent PD prevalence data. Given that PD is not an infectious or virulent disease, it is not required to be reported by health organizations such as the Centers for Disease Control and Prevention (CDC) or the World Health Organization (WHO) and little record of its prevalence remains. In 2014, the Parkinson's Foundation started the Parkinson's Prevalence Project, an initiative aimed at identifying the commonality of the disease and understanding its spread; this was later used to estimate PD prevalence [20]. After doing extensive in-depth research and communicating with some very highly regarded professionals in this field such as Dr. Connie Marras, Dr. SnehaMantri, Dr. Rodolfo Savica, and the Parkinson's Foundation, I understood that Parkinson's prevalence data had not been collected or was inaccessible to me. To solve this issue, I explored many different methods before finally setting on a proportion estimation.

Collecting the Data. In order to determine the PD prevalence in the states of North Carolina, California, Vermont, and Montana from the past four to five years, the census data giving the US population and individual state population for each year was determined. Using the earlier discussed Parkinson's Foundation Parkinson's Prevalence Project source and the 2016 US Census data, the proportion of PD cases per state population was determined. The proportion for each state was multiplied into the state populations from the past four to five years to determine the estimated amount of PD cases in each state.
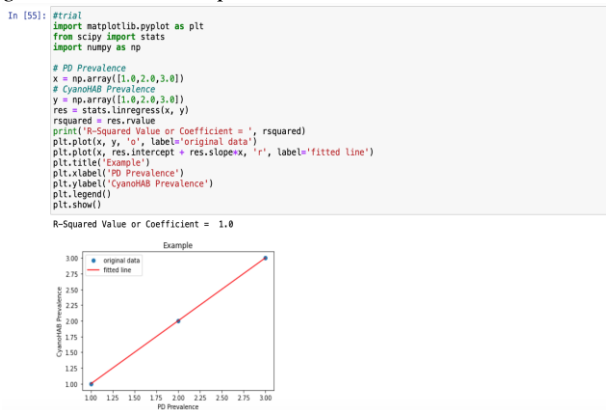
C. Conducting the Correlational Analysis

In order to determine whether Parkinson's Disease and Cyanobacterial Harmful Algal Blooms have a correlation, the collected prevalence data was analyzed using linear regression. Linear regression is a form of predictive analysis; it serves to model the relationship between two or more variables [21]. There are various forms of linear regression, most commonly including simple linear regression and multiple linear regression. Simple linear regression is performed when attempting to determine a correlation between two variables. In this case, the two variables being analyzed are CyanoHAB prevalence and PD prevalence; thus, simple linear regression was utilized in this study [22]. Simple linear regression was deemed apt for this study as it was also used in the Isenstein et al. [17] study when determining the accuracy of remote sensing to understand the water quality and characteristics of water bodies. In addition, in a comparison study carried out by Barnes and Barnes [23], it was determined that "using linear regression to estimate a trend [...] is almost always preferred over the use of [a similar unbiased linear estimator]" [23, p. 9973].

Creating the Linear Regression Python Algorithm. To determine the presence of a correlation between the PD prevalence and the CyanoHAB prevalence in each state, an algorithm that conducts linear regression and determines the r-squared coefficient was developed in the Python programming language. Python, a computer programming language similar to statistical coding language R, was utilized in this study because it has extensive statistical capabilities and is easily accessible to students. To develop the Python algorithm, the Anaconda Navigator and Jupyter Notebook were downloaded to a 2.6 GHz Dual-Core Intel Core i5 MacBook Pro using Version 10.15.6 of the MacOS Catalina. The Anaconda Navigator and Jupyter Notebook are the IDE, or the Integrated Development Environment in which the code and linear regression analysis run. After opening a Python 3 Notebook, starting in a new cell, the linear regression code was written following a similar procedure as outlined in scipy.org, a Python statistical coding reference guide. The matplotlib.pyplot, stats from scipy, math, and numpy libraries were imported into the IDE. An example of the exact code developed and its output are shown in Figure 2. The same code was used when conducting the linear regression for all four states.

https://doi.org/10.31871/IJNTR.7.7.16

International Journal of New Technology and Research (IJNTR)
ISSN: 2454-4116, Volume-7, Issue-7, July 2021 Pages 38-46

Figure 2

An Example of the Code Used to Conduct the Linear Regression and its Output

After the code to conduct the linear regression was developed, variables with the CyanoHAB prevalence for each state and each year were created. The estimated PD prevalence for all years and states were determined as detailed above and assigned to variables as well. The North Carolina PD prevalence and CyanoHAB prevalence data encoded are shown below, along with their output in Figures 3 and 4. The same code was used when assigning data to variables for all years for each state.
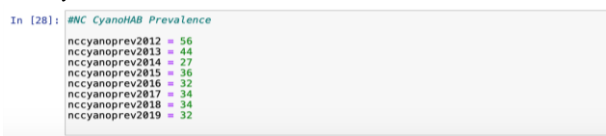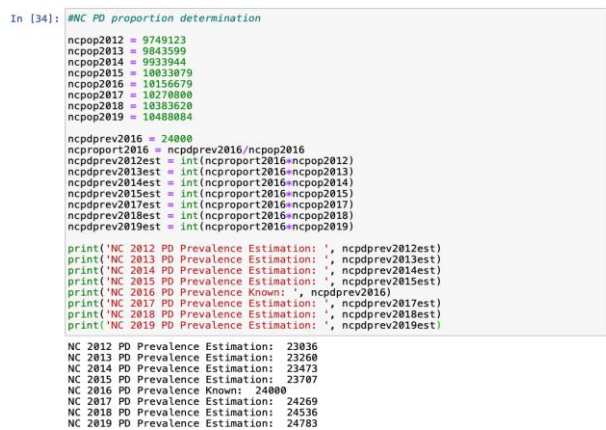
Figure 3

NC CyanoHAB Prevalence Data Encoded

Figure 4

NC PD Prevalence Data Encoded

III. STATISTICAL ANALYSIS

When analyzing the results of the linear regression, the r-squared coefficient or value is used, denoted R^2. The closer $|R^2|$ is to 1, the greater the likelihood that a correlation is present. If R^2 is closer to -1, it is likely to have a negative correlation; this signifies an inverse relationship, as one variable increases, the other decreases. If R^2 is closer to 1, it is likely to have a positive correlation; this signifies a direct relationship, as one variable increases, the other increases as well. To determine the statistical significance of the linear regression analysis, the r-squared coefficient was calculated using the Python function, stats.linregress(x,y).rvalue. The r

value was also calculated by using the math.sqrt() function to take the square root of the determined r-squared coefficient.

IV. METHOD JUSTIFICATION

While this study itself does not directly use satellite imagery to determine the presence of CyanoHABs around the US (satellite maps were used to count CyanoHAB prevalence), remote sensing is a common method that has been utilized by various peer-reviewed studies to effectively and efficiently conduct research projects. "Monitoring Variation in Sea Surface Temperature in the Nakdong River Estuary, Korea, Using Multiple Satellite Images," conducted by Eom et al. [24], measured the temperature of the surface of the sea in Korea specifically utilizing satellite imaging and remote sensing [24]. They also conducted a linear analysis and found a line of best fit for their data, as well as used r-squared coefficient values for their results. The use of remote sensing and linear analysis in this study shows that this is a viable and reliable method that can be replicated in similar studies such as this one. In the study conducted by Isenstein et al. [17], linear regression analysis in the programming language R was performed; given that this method was verified, this study too performed linear regression in a similar, more accessible programming language, Python. In order to prevent bias and error in the code, the code was modeled directly after the linear regression example code provided by the Python official reference guide for statistical functions. These reference guides, specifically SciPy for this project, are produced directly by the creators of the Python language, and they give detailed, objective information on how to write a functional statistical analysis code.

V. DATA

Table 1

This table contains the CyanoHAB and PD Prevalence determined for each of the four states.

	NC		MT		VT		CA	
	CyanoHAB Prevalence	PD Prevalence	CyanoHAB Prevalence	PD Prevalence	CyanoHAB Prevalence	PD Prevalence	CyanoHAB Prevalence	PD Prevalence
2014	27	23473	----	----	----	----	----	----
2015	36	23707	----	----	217	1804	----	----
2016	36	24000	----	----	148	1800	56	85100
2017	37	24269	33	3035	268	1802	114	85512
2018	34	24536	56	3061	277	1807	160	85855
2019	----	----	21	3080	----	----	196	85757
2020	----	----	64	3112	----	----	311	86339

Figure 5

This figure contains the linear regressions graphs for all states

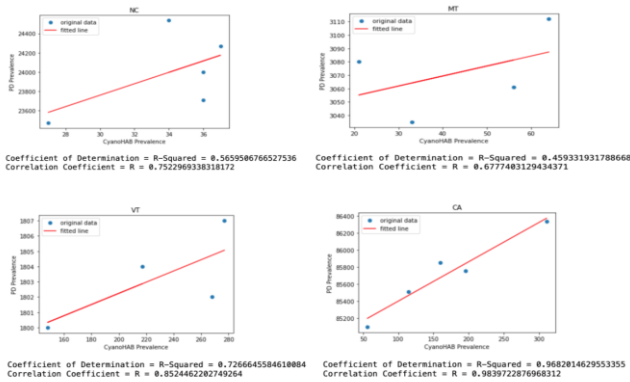**Table 2**

This table contains the acquired r-squared coefficient and r-value for each state.

	R-Squared Coefficient	R-Value
NC	0.57	0.75
MT	0.46	0.68
VT	0.73	0.85
CA	0.97	0.98

VI. RESULTS

The r-squared coefficient and r-values given in Table 2 will be used to determine the presence and magnitude of a correlation between the variables of the PD prevalence and CyanoHAB prevalence. According to Keith G Calkins [25], an Associate Professor of Math and Science at Andrews University, intervals given below in Table 3 determine the strength of a correlation.

Table 3

Table contains the R-Squared magnitude intervals and the corresponding correlation levels

|R-Squared Interval|	Correlation Level
0.9-1.0	Very High Correlation
0.7-0.9	High Correlation
0.3-0.7	Moderate Correlation
<0.3	Little or No Correlation

Table 4

Table contains the R-Squared Coefficient and corresponding correlational level

	R-Squared Coefficient	Correlational Level
NC	0.57	Moderate Correlation
MT	0.46	Moderate Correlation
VT	0.73	High Correlation
CA	0.97	Very High Correlation

Based on Tables 3 and 4, it appears that all states had a moderate to high correlation between CyanoHABs and PD.

VII. DISCUSSION

A. Implications

This study has various implications for Parkinson's Disease and aging, as well as climate change and its effect on humans. How PD is induced in the body has eluded scientists, but they have focused upon genetic predispositions and genes such as SNCA, PINK1, and LRRK2, as well as pesticides like Paraquat and Rotenone, all of which have been linked to Parkinsonism [26]. As discussed previously in the literature review, studies by Arif et al. [11] and Cox et. al [10] have thoroughly examined BMAA and its function in causing PDC in the body. My study goes one step further to correlate PD to CyanoHABs, BMAA's reservoir and even climate change. As explained by Anneville et al. [2], "Cyanobacteria [...] are [...] well adapted to the environmental conditions associated with global warming, such as high nutrient loading, rising temperatures, [...] and salinization" [2, p. 442]. Given that cyanobacteria and CyanoHABs are results of climate change according to Anneville et al. [2] and Glibert & Burford [15], PD may be an additional consequence of climate change. Climate change has long been suspected of having negative consequences not only on the environment and animal habitats but also on human health: "Weather and climate factors are known to affect the transmission of [...] diseases [and] chemicals around the environment. Climate change [has] important impacts on the dispersion of pathogens and chemicals in the environment" [27, p. 508].

In addition, the population of many developed nations is aging quickly. According to one report, 25% of the US population is expected to be older than 65 by 2050 [3]. Given this, according to Ortman et al. [3], the US is actually considered younger than many other nations: over 20% of the population in Germany, Italy, and Japan were older than 65 in 2012. The likelihood of developing neurodegenerative disease such as PD increases greatly for older individuals, especially men; thus with an aging population, more people are likely to develop PD and more research such as this is necessary and helpful to gain insight into an uncharted

disease and its potential causes [20],[28]. Given that PD has been positively correlated with CyanoHABs, scientists can investigate mechanisms to limit human interaction with CyanoHABs or look into CyanoHAB prevention possibilities such as limiting fertilizer use, filtering runoff, and preventing global warming.

B. Limitations

While a correlation between CyanoHABs and PD was found in the four states studied, a few limitations were encountered when performing this study. One major limitation experienced was a lack of data available in this field. I had great difficulty finding both PD and CyanoHAB prevalence data. Because CyanoHABs are a relatively newly-investigated phenomena and climate awareness is still in its beginning stages, little data was available for a retrospective study such as this one, and it greatly limited the amount of regions I was able to study in this project. While there were satellite images that were available for the US, these images were difficult to decipher and analyze, as well as difficult to procure as the NASA and similar websites require various authorizations to access.

Another limitation faced in this study was the unavailability of PD prevalence data. When initially looking for PD prevalence data to utilize for this project, I was able to find a few different prevalence numbers for states and countries such as Canada, Hawaii, and California, and I assumed with more extensive research that greater amounts of data would be available. However, because Parkinson's Disease is not a transmissible or dangerous disease, health organizations such as the WHO and CDC don't record its prevalence. To overcome this challenge, a proportion developed using the previously mentioned 2016 PPP study and US Census data was utilized to estimate PD prevalence numbers. While moderate to high correlations between CyanoHABs and PD prevalence were found in this study, estimated PD prevalence numbers were used; this was another limitation of this study. Because correlational analysis and data was only collected for four states, conclusions can only be made specifically for North Carolina, Montana, Vermont, and California, however, we can effectively infer that states in similar locations and circumstances will have similar correlations.

C. Future Direction & Conclusions

Increase Data Accuracy. One of the largest limitations of this study was the unavailability of both CyanoHAB and PD data. Thus, fewer states were studied for a shorter period of time that initially expected and estimated PD prevalence data was utilized. In the future, the remaining 46 states' PD and CyanoHAB correlation could be studied if/when more data becomes available. To overcome the lack of data in the future without the constraints of time, resources, and the ongoing pandemic, I or other researchers could reach out to the largest medical centers of each state to get more accurate PD prevalence numbers. This would allow the conducted analyses to be more accurate and indicative of the future of PD. Similarly, other countries that keep better track of their PD prevalence can also be studied. For example, one Portuguese study conducted a PD prevalence project in 1994,

and a similar method of reaching out to various doctors to determine PD prevalence can be used again to determine Portuguese and other countries' PD prevalence [29].

To get more accurate CyanoHAB prevalence data, remote sensing and satellite use could prove helpful; according to Victor Klemas, Professor of Marine Studies at the School of Marine Science and Policy at the University of Delaware, "Satellites are at the forefront of ocean-color mapping because the magnitude and variability of chlorophyll concentrations are poorly known on global and regional scales, largely due to high spatial and temporal variability of marine phytoplankton" [30, p. 36]. Another study by DJ Grimes et al. [31] corroborates the benefits of remote sensing and satellite imaging; specifically, this study focuses on how satellite imaging has been used to determine the marine algae concentrations present in different regions of water bodies. This is another future direction that can be pursued, different algae and cyanobacteria concentrations in CyanoHABs can be studied through satellite imagery and can further be correlated with PD to better understand how different concentrations of algae contribute to disease.

Changing the Study Design. Earlier, a paper by Dr. Dean R Hess [16] was referenced when describing the retrospective study format used in this project. While Hess's paper concludes that prospective studies are preferable to retrospective studies, given that prospective studies require collection of data in future years rather than using previously collected data; thus, a prospective study is not feasible for this study. Hess does admit that retrospective studies can be beneficial in that they can give useful information about and "can help to focus the study question, clarify the hypothesis, determine an appropriate sample size, and identify feasibility issues for a prospective study" [16, p. 1171]. Thus, seeing that a retrospective study such as the one performed for this project is helpful anticipatory research for prospective studies, a prospective study following current PD and control subjects and water bodies near their locations could be pursued in the future.

Further Statistical Analyses. This study's statistical analysis is limited to the conduction of a linear regression test as well as a correlation and determination coefficient analysis. In the future, further statistical analyses such as the ANOVA test and Tukey Kramer test can be done to further test the strength of the correlation between CyanoHABs and PD determined by this study. For example, in the ANOVA test, also known as the Analysis of Variance test, the means of the groups can be determined and compared in the Tukey Kramer test as long as the ANOVA's F-value is significant and the null hypothesis is rejected [32]. This would allow a comparison of the similarity of the correlations between states; thus, even if a state did not have a high correlation coefficient, if it was proven similar to those that do have a large r-squared value, it could help develop a more unambiguous conclusion and comment further on the correlation between CyanoHABs and PD.

While this study has supported a correlation between CyanoHABs and PD, we cannot assume that this is also indicative of causation between the two variables. Another future direction that can be pursued is studying this data under

Bradford Hill's Criteria of Causation. The nine criteria include strength of association, consistency, specificity, temporality, biological gradient, plausibility, coherence, experiment, and analogy [33]. Each of these criteria look at a different aspect of association from which one can then infer causation. For example, strength of association looks at the fact that the larger an association is between 2 variables, the more likely it is that the association is causal [33]. Similarly, consistency and specificity look at the fact that the association remains consistent even among changing factors and that the more specific an association is, the more likely it is that it is causal. Thus, specificity would be a factor that has already been supported by this study, as a one-to-one relationship is being studied, how CyanoHABs and PD are related [33].

Conclusions & Context. The hypothesis that there was a correlation between CyanoHABs and PD has been supported by the results and data. Given that the correlation coefficients found were positive in all four states, a positive correlation between CyanoHABs and PD has been found.Thus, the answer to my research question is that Cyanobacterial Harmful Algal Blooms and Parkinson's Disease prevalence rates have positively affected each other. While in this study no specific dependent or independent variables have been defined (as correlation not causation is being studied), the CyanoHAB prevalence was treated as the independent variable and the PD prevalence was treated as the dependent variable. This was done because, if there was not only correlation but also causation between the 2 variables, CyanoHABs would result in an increase in PD prevalence, not vice versa. Because CyanoHABs are a result of global warming and climate change, PD, along with cholera, ebola, and tuberculosis, may be another disease linked to climate change [34].

REFERENCES

[1] Cox, P., Banack, S., Murch, S., Rasmussen, U., Tien, G., Bidigare, R., . . . Bowers, W. (2005). Diverse Taxa of Cyanobacteria Produce β-N-Methylamino-L-Alanine, a Neurotoxic Amino Acid. Proceedings of the National Academy of Sciences of the United States of America, 102(14), 5074-5078. Retrieved November 17, 2020, from http://www.jstor.org/stable/3375181

[2] Anneville, O., Domaizon, I., Kerimoglu, O., Rimet, F., &Jacquet, S. (2015). Blue-Green Algae in a "Greenhouse Century"? New Insights from Field Data on Climate Change Impacts on Cyanobacteria Abundance. Ecosystems, 18(3), 441-458. Retrieved October 27, 2020, from http://www.jstor.org/stable/43677726

[3] Ortman, J. M., Velkoff, V. A., & Hogan, H. (2014, May). An Aging Nation: The Older Population in the United States. Time. https://time.com/wp-content/uploads/2015/01/p25-1140.pdf

[4] Clarke, C. (2007). Parkinson's Disease. BMJ: British Medical Journal, 335(7617), 441-445. Retrieved September 25, 2020, from http://www.jstor.org/stable/20507607

[5] Atkins, G. (2016). Aging in America: An Agenda for an Era of New Possibilities. Generations: Journal of the American Society on Aging, 40(4), 6-8. doi:10.2307/26556240

[6] Berdalet, E., Fleming, L. E., Gowen, R., Davidson, K., Hess, P., Backer, L. C., Moore, S. K., Hoagland, P., & Enevoldsen, H. (2015). Marine harmful algal blooms, human health and wellbeing: challenges and opportunities in the 21st century. Journal of the Marine Biological Association of the United Kingdom. Marine Biological Association of the United Kingdom, 2015, 10.1017/S0025315415001733. https://doi.org/10.1017/S0025315415001733

[7] Mishra, S., Stumpf, R.P., Schaeffer, B.A. et al. Measurement of Cyanobacterial Bloom Magnitude using Satellite Remote Sensing. Sci Rep 9, 18310 (2019). https://doi.org/10.1038/s41598-019-54453-y

[8] Paerl, H., &Whitall, D. (1999). Anthropogenically-Derived Atmospheric Nitrogen Deposition, Marine Eutrophication and Harmful Algal Bloom Expansion: Is There a Link? Ambio, 28(4), 307-311. Retrieved November 17, 2020, from http://www.jstor.org/stable/4314901

[9] Berdalet, E., Tester, P., Chinain, M., Fraga, S., Lemée, R., Litaker, W., . . . Zingone, A. (2017). Harmful Algal Blooms in Benthic Systems: Recent Progress and Future Research. Oceanography, 30(1), 36-45. Retrieved November 3, 2020, from http://www.jstor.org/stable/24897840

[10] Cox, P., Davis, D., Mash, D., Metcalf, J., &Banack, S. (2016). Dietary exposure to an environmental toxin triggers neurofibrillary tangles and amyloid deposits in the brain. Proceedings: Biological Sciences, 283(1823), 1-10. Retrieved October 11, 2020, from http://www.jstor.org/stable/24762478

[11] Arif, M., Kazim, S., Grundke-Iqbal, I., Garruto, R., &Iqbal, K. (2014). Tau pathology involves protein phosphatase 2A in Parkinsonism-dementia of Guam. Proceedings of the National Academy of Sciences of the United States of America, 111(3), 1144-1149. Retrieved October 11, 2020, from http://www.jstor.org/stable/23770442

[12] Hansen, H. (2020, April 2). Fallacies (Stanford Encyclopedia of Philosophy). Stanford Encyclopedia of Philosophy. https://plato.stanford.edu/entries/fallacies/

[13] Nunes-Costa D, Magalhães JD, G-Fernandes M, Cardoso SM and Empadinhas N (2020) Microbial BMAA and the Pathway for Parkinson's Disease Neurodegeneration. Front. Aging Neurosci. 12:26. doi: 10.3389/fnagi.2020.00026

[14] [14] Dobbins, J., Solomon, R., Chase, M., Henry, R., Larrabee, F., Lempert, R., . . . Shatz, H. (2015). Climate Change. In Choices for America in a Turbulent World: Strategic Rethink (pp. 69-84). RAND Corporation. Retrieved December 27, 2020, from http://www.jstor.org/stable/10.7249/j.ctt17mvhfj.13

[15] Glibert, P., & Burford, M. (2017). Globally Changing Nutrient Loads and Harmful Algal Blooms: Recent Advances, New Paradigms, and Continuing Challenges. Oceanography, 30(1), 58-69. Retrieved October 27, 2020, from http://www.jstor.org/stable/24897842

[16] Hess, D. R. (2004, October). Retrospective studies and chart reviews. PubMed. https://pubmed.ncbi.nlm.nih.gov/15447798/

[17] Isenstein, E., Trescott, A., & Park, M. (2014). Multispectral Remote Sensing of Harmful Algal Blooms in Lake Champlain, USA. Water Environment Research, 86(12), 2271-2278. Retrieved November 18, 2020, from http://www.jstor.org/stable/24585738

[18] Cyanobacteria (Blue-Green Algae) Tracker. (2020, July 1). Vermont Department of Health. https://www.healthvermont.gov/tracking/cyanobacteria-tracker

[19] Surface Water - Freshwater Harmful Algal Blooms - Surface Water – Freshwater Harmful Algal Blooms - California Open Data. (2020). California Surface Water – Freshwater Harmful Algal Blooms. https://data.ca.gov/dataset/surface-water-freshwater-harmful-algal-blooms/resource/c6f760be-b94f-495e-aa91-2d8e6f426e11

[20] Statistics. (2020). Parkinson's Foundation. https://www.parkinson.org/Understanding-Parkinsons/Statistics?utm_source=google&utm_medium=adgrant&utm_campaign=&utm_term=&gclid=CjwKCAjw5p_8BRBUEiwAPpJO66d0idixsJdt8ki2DHQc0ykKLX1KM8FQArhgBZLWQK_c4qR9I71DfhoCNPUQAvD_BwE

[21] Linear Regression. (1997). Linear Regression. http://www.stat.yale.edu/Courses/1997-98/101/linreg.htm

[22] Klotz, J. (1995). UPDATING SIMPLE LINEAR REGRESSION. StatisticaSinica, 5(1), 399-403. Retrieved December 1, 2020, from http://www.jstor.org/stable/24305577

[23] Barnes, E., & Barnes, R. (2015). Estimating Linear Trends: Simple Linear Regression versus Epoch Differences. Journal of Climate, 28(24), 9969-9976. Retrieved October 30, 2020, from http://www.jstor.org/stable/26195794

[24] Eom, J., Park, W., Syifa, M., Lee, C., & Yoon, S. (2019). Monitoring Variation in Sea Surface Temperature in the Nakdong River Estuary, Korea, Using Multiple Satellite Images. Journal of Coastal Research, 183-189. Retrieved November 4, 2020, from https://www.jstor.org/stable/26778952

[25] Calkins, K. G. (2005). Correlation Coefficients. Correlation Coefficients Andrews University. https://www.andrews.edu/%7Ecalkins/math/edrm611/edrm05.htm

[26] Brown, T., Rumsby, P., Capleton, A., Rushton, L., & Levy, L. (2006). Pesticides and Parkinson's Disease-Is There a Link? Environmental Health Perspectives, 114(2), 156-164. Retrieved November 27, 2020, from http://www.jstor.org/stable/3436503

[27] Boxall, A., Hardy, A., Beulke, S., Boucard, T., Burgin, L., Falloon, P., . . . Williams, R. (2009). Impacts of Climate Change on Indirect Human Exposure to Pathogens and Chemicals from Agriculture.

https://doi.org/10.31871/IJNTR.7.7.16

International Journal of New Technology and Research (IJNTR)
ISSN: 2454-4116, Volume-7, Issue-7, July 2021 Pages 38-46

Environmental Health Perspectives, 117(4), 508-514. Retrieved November 15, 2020, from http://www.jstor.org/stable/25478152

[28] Marras, C., Beck, J. C., Bower, J. H., Roberts, E., Ritz, B., Ross, G. W., Abbott, R. D., Savica, R., Eeden, D. V. S. K., Willis, A. W., & Tanner, C. M. (2018, July 10). Prevalence of Parkinson's disease across North America. Npj Parkinson's Disease. https://www.nature.com/articles/s41531-018-0058-0?error=cookies_n ot_supported&code=dd643e5c-c111-4045-b2c4-d7892c766ae7

[29] Dias, J., Felgueiras, M., Sanchez, J., Gonçalves, J., Falcäo, J., & Pimenta, Z. (1994). The Prevalence of Parkinson's Disease in Portugal: A Population Approach. European Journal of Epidemiology, 10(6), 763-767. Retrieved November 14, 2020, from http://www.jstor.org/stable/3581715

[30] Klemas, V. (2012). Remote Sensing of Algal Blooms: An Overview with Case Studies. Journal of Coastal Research, 28(1A), 34-43. Retrieved October 20, 2020, from http://www.jstor.org/stable/41332047

[31] Grimes, D., Ford, T., Colwell, R., Baker-Austin, C., Martinez-Urtaza, J., Subramaniam, A., & Capone, D. (2014). Viewing Marine Bacteria, Their Activity and Response to Environmental Drivers from Orbit: Satellite Remote Sensing of Bacteria. Microbial Ecology, 67(3), 489-500. Retrieved October 20, 2020, from http://www.jstor.org/stable/24542326

[32] Glen, S. (2021, January 1). Tukey Test / Tukey Procedure / Honest Significant Difference. Statistics How To. https://www.statisticshowto.com/tukey-test-honest-significant-differen ce/

[33] Fedak, K., Bernal, A., & Gross, S. (2015). Applying the Bradford Hill criteria in the 21st century: how data integration has changed causal inference in molecular epidemiology. PubMed Central (PMC). https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4589117/

[34] Biello, D., &Biello, D. (2008, October 8). Deadly by the Dozen: 12 Diseases Climate Change May Worsen. Scientific American. https://www.scientificamerican.com/article/twelve-diseases-climate-c hange-may-make-worse/#:%7E:text=Bird%20flu%2C%20cholera%2 C%20Ebola%2C,Wildlife%20Conservation%20Society%20(WCS)