The relationship between maritime phosphate pollution and socioeconomic wellbeing

10.17975/sfj-2021-012

Tony Hu[1], William Zhou[1], Andrew Li[1], and Dhananjay Patki[1]

[1] Westmount Charter School, Alberta, Canada

ABSTRACT: Phosphate rock reserves are expected to deplete in the next 50-100 years, with the point of highest phosphorus production predicted to be in 2030. Phosphate, the base of many fertilizers, is a non-renewable resource. Ocean phosphate concentrations provide a good indication of global fertilizer use, since agricultural runoff often contributes to increases in ocean phosphate concentration. This study explores the relationship between the concentration of phosphate in a nation's maritime borders and the nation's score on the Social Progress Index. The study aims to link findings with possible approaches to help meet two of the United Nations' (UN) Sustainable Development Goals: creating sustainable communities, and conserving and sustainably using oceans. Phosphate concentration data were acquired from the National Oceanic and Atmospheric Administration and compared with factors of social welfare from the 2018 Social Progress Index. For each nation on the Social Progress Index, the nation's score on every factor was separately compared to ocean phosphate concentration data within that nation's maritime borders, and a linear regression was performed for each comparison. The results indicate countries ranking higher on the Social Progress Index generally have greater ocean phosphate concentrations, suggesting that countries of higher social welfare contribute more to global phosphate use or have greater amounts of fertilizer runoff. The findings should be considered by developed nations to inform decisions around pollution reduction as well as developing nations aiming for sustainable social progress. Both should consider the environmental effects that social progress has or will have on the greater global community, of which the significance to sustainable resource development and environmental protection is invaluable.

KEYWORDS: development, ocean, agriculture, pollution

INTRODUCTION

Ocean phosphate levels have experienced a global increase, largely explained by the growing use of phosphate-based fertilizer. Fertilizers were first commercialized in 1910, and are commonplace in modern agriculture [1]. This is evidenced by the increase of crop yield globally, with the average wheat yield of the United Kingdom exceeding eight tons per hectare of land [2]. Further, increased fertilizer use has heightened the problem of runoff, where excess nutrients are dispersed into the ecosystem and catalyze the growth of harmful bacteria and algae [3]. As such, the increase in fertilizer use poses a concern in the sustainable use and development of natural resources.

Current research generally supports the notion that agricultural productivity is highly important to economic development [4-5]. Increased fertilizer runoff is indicative of this relationship. However, we wish to challenge this long-held belief because of the introduction of a competing effect in modern agriculture, that of pollution control. We predict that stricter water pollution control policies in the past several years, such as those in the Netherlands [6], could reverse the relationship between fertilizer runoff and social development.

The Social Progress Index (SPI) contains scorecards between 2014 and 2019. The 2019 SPI contains scorecards from 238 countries that were scored based on 63 factors of social welfare. Factor scores given to countries were based on various sources (e.g., Food and Agricultural Organization of the United Nations, the World Bank, Gallup World Poll), and averaged to produce a country's overall score. The countries are ranked based on their overall scores, and their scorecards are published on the SPI [7].

The aim of this study is to investigate the effects of humanity's long-term use of fertilizers and contextualize it within the UN Sustainable Development Goals on making cities inclusive, safe, resilient and sustainable; and conserving and sustainably using the oceans, seas and marine resources [8]. This can be assessed by determining the correlation between the phosphate concentration within a nation's maritime boundary and 63 factors of social welfare in the nation. We hypothesize there to be a negative relationship between a country's score on the Social Progress Index and the country's ocean phosphate concentration.

METHODS

The global data used in this study comes from two major databases. Ocean phosphate concentration data was taken from the National Oceanic and Atmospheric Administration's (NOAA) 2018 World Ocean Atlas project, and contains data at multiple depths for all regions of the world [9]. The social welfare data was taken from the SPI for 2018, and contains data on factors of social welfare for 238 geographic areas [7]. The phosphate dataset contained phosphate concentration at each depth, but many values were missing for deeper depths, as seen in Figure 1.

Existing 10 m concentration values were used for the analysis when available. If absent, a fifth degree polynomial line of best fit model was used to predict the phosphate levels at the 10 m depth based on the data from other oceanic depths (Figure 2). If less than 10 data points were available at the other depths, the data were discarded, as the model is unable to make meaningful predictions with limited data. Predictions which exceeded 1.5 times the interquartile range from the nearest quartile of all data values were removed. The final distribution of all phosphate concentrations at the 10 m depth is shown in Figure 3.

Next, each phosphate data point was mapped to a region in the SPI using its geographic coordinates. Oceanic boundaries for the region were determined by the nation's Exclusive Economic Zone (EEZ), which was obtained from the EEZ dataset [10]. The 2019 world EEZ boundaries were used, which is shown in Figure 4. A world map with every location's phosphate concentration overlaid is shown in Figure 5.

Using the world EEZ map, the provided longitudes and latitudes were associated with the nation they belonged to. For example, in order to acquire data for the United States (US), all geographical coordinates within the US' EEZ were gathered, and a mean was estimated for the corresponding

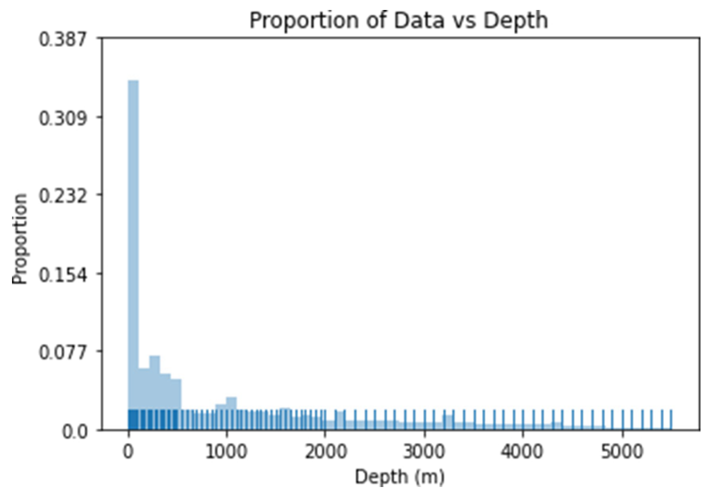
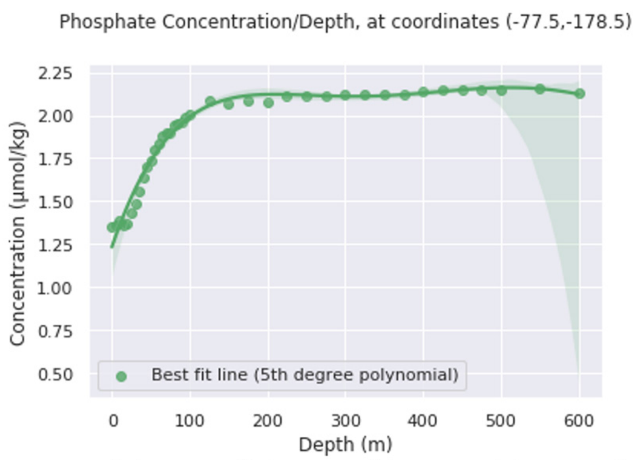


Figure 1. Proportion of total dataset vs. depth. Shown is a histogram of the proportion of the total dataset corresponding to each depth.



$$y = 1.56 \times 10^{-14}x^5 - 3.13 \times 10^{-10}x^4 + 2.36 \times 10^{-7}x^3 - 8.32 \times 10^{-5}x^2 + 1.36 \times 10^{-2}x + 1.26$$

Figure 2. An example of phosphate concentration by ocean depth. The model was created using a fifth degree polynomial line of best fit; $R^2 = 0.983$.

phosphate concentrations. This mean concentration and the nation's scores for each SPI factor was plotted to determine a visual relationship. Subsequently, the strength of the relationship between mean phosphate levels and SPI factors were determined using Pearson correlation.

Attempts were made to minimize error as much as possible. Within the Social Progress Index, some scores were empty. Instead of discarding rows with missing data, gaps were filled with the median score of the category. Individually, each missing value is unlikely to be very far from the category median because the factors in each category are related, making the fill-in relatively accurate.

Google Colaboratory (Google LLC, Mountain View,

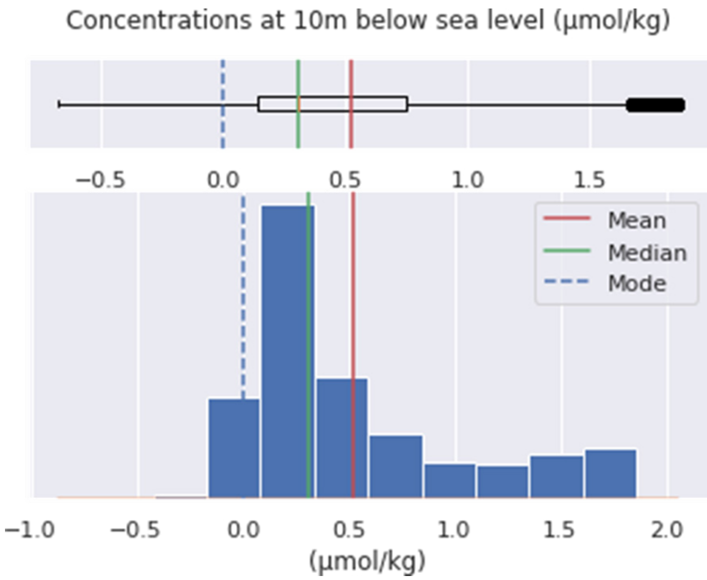


Figure 3. Distribution of all phosphate concentrations at 10 m depth. Shown is a boxplot and histogram of all phosphate concentration values after outlier removal.

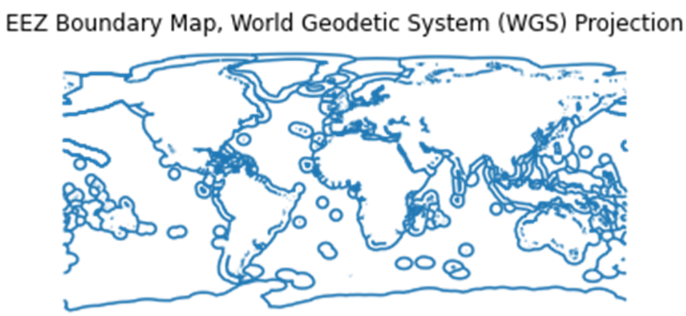


Figure 4. Exclusive Economic Zone boundaries from 2019.

California, US) and Python 3 (Python Software Foundation, Delaware, US) were used for analysis. The Python libraries Numpy, Pandas, Matplotlib, Scipy, Tqdm, Shapely, Geopandas, Fiona, Rtree, Cartopy, Tqdm, and Seaborn were used in the creation of the analysis and visualization programs. Extra installed pip-packages include Cartopy, Geopandas, Pycountry, Convert, and Countryinfo. All programs were originally created and coded manually. The full data analysis is available upon request.

RESULTS

A total of 23,660 phosphate data points were analyzed, of which 11,207 fell into the EEZ of at least one geographic region. After performing regressions across the SPI factors globally, only weak and non-linear relationships were found. In consideration of variations within the global data, another set of linear regressions was performed with the data disaggregated by region. The new data produced correlation coefficients which were greater in magnitude, with some ranging from 0.6 to 0.7 (Tables 1-2).

To verify the correct use of linear regression, three assumptions must be met: the variables are quantitative, the relationship is roughly linear, and the relationship is homoscedastic. All variables on the SPI and ocean phosphate concentration were quantitative. Residual plots for each of the 378 factor-region pairs were plotted and checked for possible patterns to verify the assumption of linearity. These residual plots were then used to verify the assumption of homoscedasticity. Most of the scatter plots met these two assumptions and the ones that failed did not have significant correlation coefficients and were not further analyzed. The scatter plots yielding the three greatest magnitude correlations are shown with their respective residual plots in Figures 6-8. The scatterplot of SPI vs. ocean phosphate

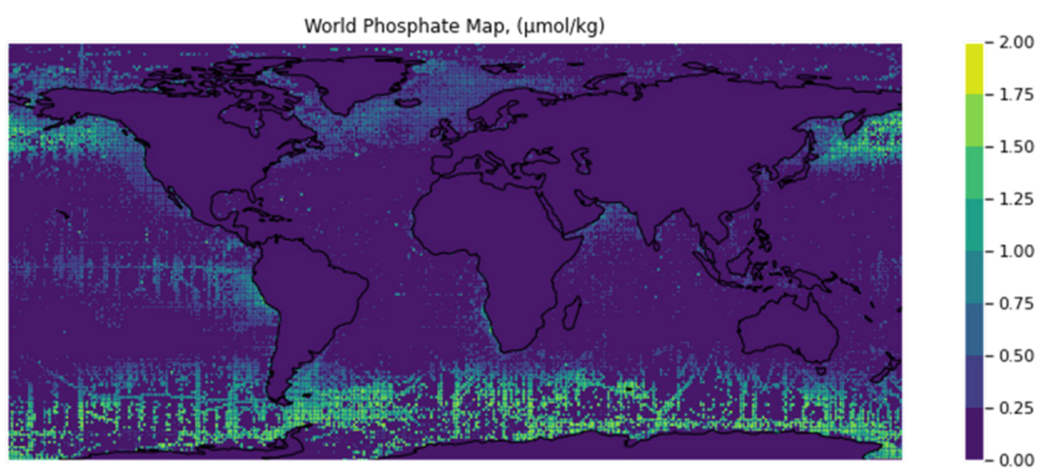


Figure 5. World map with phosphate concentration overlay showing the phosphate concentration at 10 m depth.

STEM Fellowship Journal

Table 1. Social factors of greatest positive correlation with phosphate concentration. The factor-region pairs that produced the greatest positive correlation coefficient with phosphate concentration.

Factor	Correlation coefficient (r)
Quality of electricity supply (1=low; 7=high), North America	0.625
Years of tertiary schooling, South America	0.702
Traffic deaths (deaths/100,000 people), Asia	0.617
Greenhouse gas emissions (CO_2 equivalents/GDP), Africa	0.464

Table 2. Social factors of greatest negative correlation with phosphate concentration. The factor-region pairs that produced the greatest negative correlation coefficient with phosphate concentration.

Factor	Correlation coefficient (r)
Gender parity in secondary enrollment (distance from parity), North America	-0.619
Homicide rate (deaths/100K people), South America	-0.655
Years of tertiary schooling, Asia	-0.435
Access to basic knowledge, Africa	-0.429

concentration may be of interest and is shown in Figure 9. However, it failed to meet the homoscedasticity assumption.

DISCUSSION

The largest correlation was 0.702, between years of tertiary schooling and ocean phosphate concentration in South America. There is a moderately strong, positive, linear relationship between the two variables. The strength of the national economy may be associated with both variables, as it would likely have a positive correlation with both factors. A country with a stronger economy may have a greater proportion of students enrolled in tertiary schooling for a longer period and would also likely have greater amounts of phosphate runoff from increased production and imports. The factor with the greatest negative correlation was homicides per 100,000 people in South America. The correlation coefficient was -0.655, shown in Figure 8 as a moderately strong negative linear relationship. This suggests South American nations with low homicide rates had high levels of phosphate in their waters. Nations with strong economies tend to have more robust law enforcement, and have more advanced agricultural sectors with higher use of fertilizers.

Squaring the two correlation coefficients used previously yields an R^2 of 0.493 and 0.429 respectively. It can be concluded that 49.3% of the variation in years of tertiary schooling in South America can be explained by the linear relationship with ocean phosphate concentration, and 42.9% of the variation in homicides per 100,000 people can be explained by the linear relationship with ocean phosphate concentration. These statements do not imply that changes in the explanatory variables cause changes in the response variables.

North and South America had the largest correlation coefficients, and this pattern was consistent for most factors analyzed. A possible explanation for North America's highly consistent correlation is the relative economic stability in the region. Many of the SPI factors are closely related to economic factors, so if there were a general association between economic strength and ocean phosphate concentration for most countries, North America would have consistently large correlations across those factors.

Since the data represented in Figure 9 did not meet the assumptions of linear regression, no conclusions can be drawn regarding the overarching relationship of socioeconomic wellbeing and ocean phosphate concentration. While specific socioeconomic factors are positively or negatively correlated with ocean phosphate pollution, overall socioeconomic wellbeing may not be correlated with ocean phosphate pollution.

In general, current research agrees with the notion that there is a positive relationship between agricultural productivity and economic growth, but there has been little research into the causality [11]. However, this does not explain the moderately strong negative correlation between access to basic knowledge in Africa and ocean phosphate concentration found in the present study. Reimers and Klasen state that "the literature so far can be judged as rather inconclusive about the role of education for agricultural productivity in the international context", with various studies

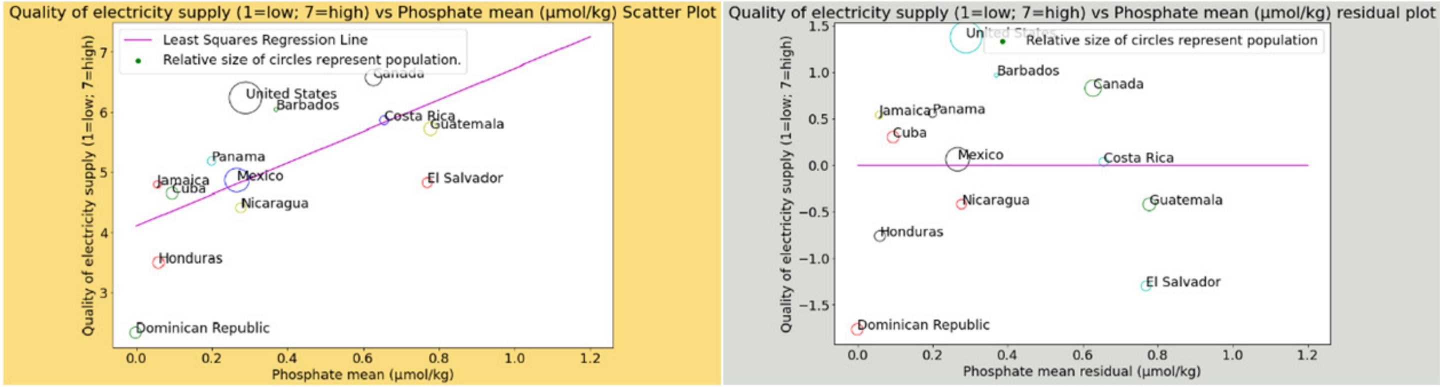


Figure 6. Quality of electricity supply vs. phosphate concentration in North America shown as a scatter plot (left) and residual plot (right).

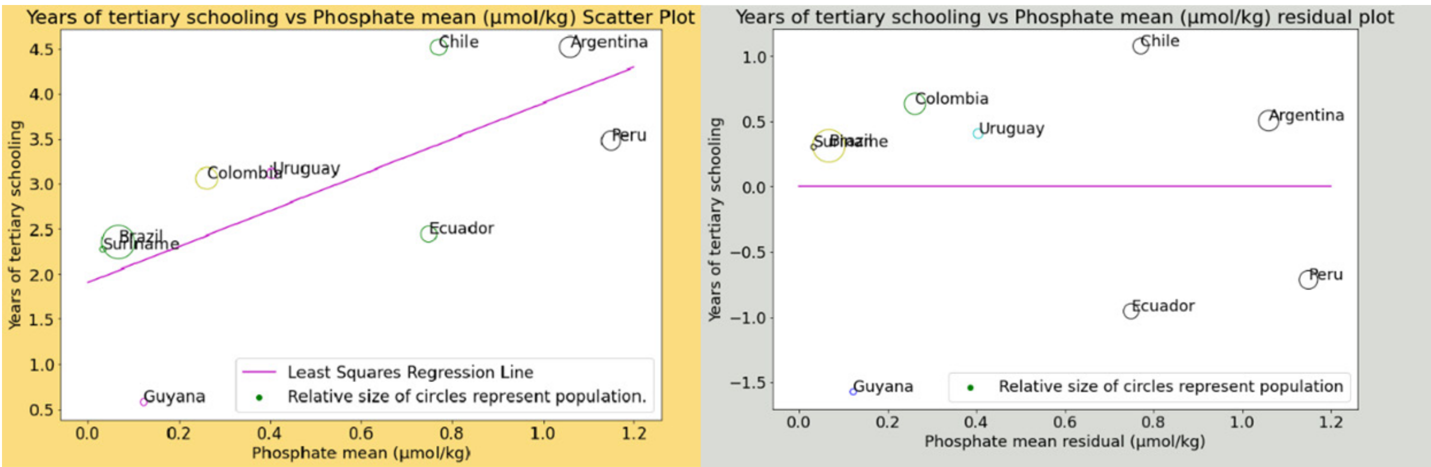


Figure 7. Years of tertiary schooling vs. phosphate concentration in South America shown as a scatter plot (left) and residual plot (right).

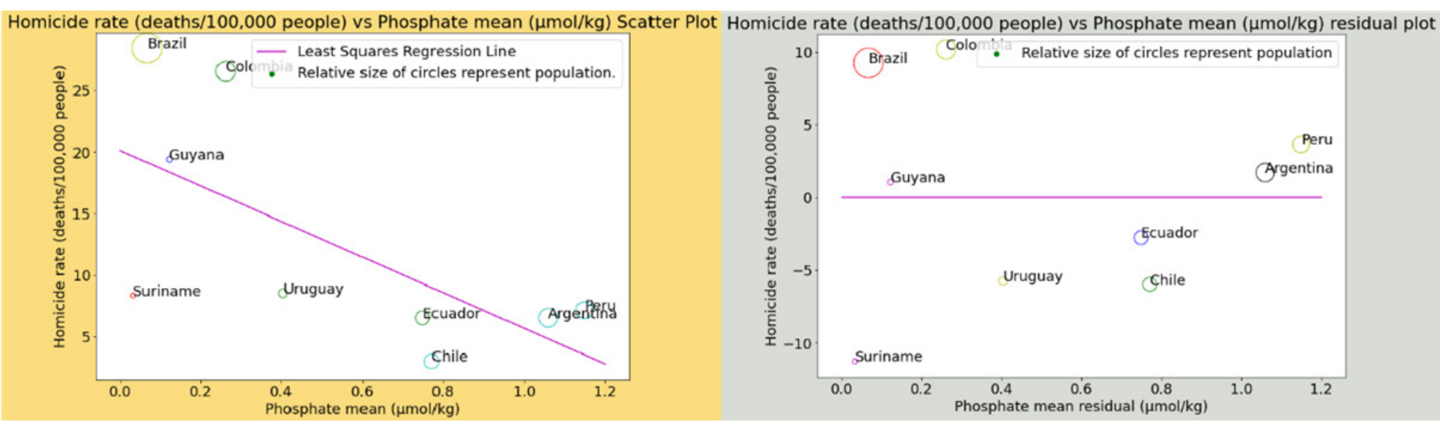


Figure 8. Homicide rate vs. phosphate concentration in South America shown as a scatter plot (left) and residual plot (right).

obtaining positive, negative, and no correlations between the two [12]. They assert that previous studies "fail to detect the expected impact because they are using inadequate variables (enrolment and literacy rate) to approximate the stock of education" and advocate the use of the education dataset compiled by Barro and Lee [12-13]. This dataset is now widely accepted in economics, and we acknowledge the work of others such as Reimers and Klasen are more accurate regarding the specific relationship between education and agricultural productivity [12]. Thus, the negative relationship obtained in this study may not suggest all African countries exhibit lower agricultural productivity with increased access

STEM Fellowship Journal

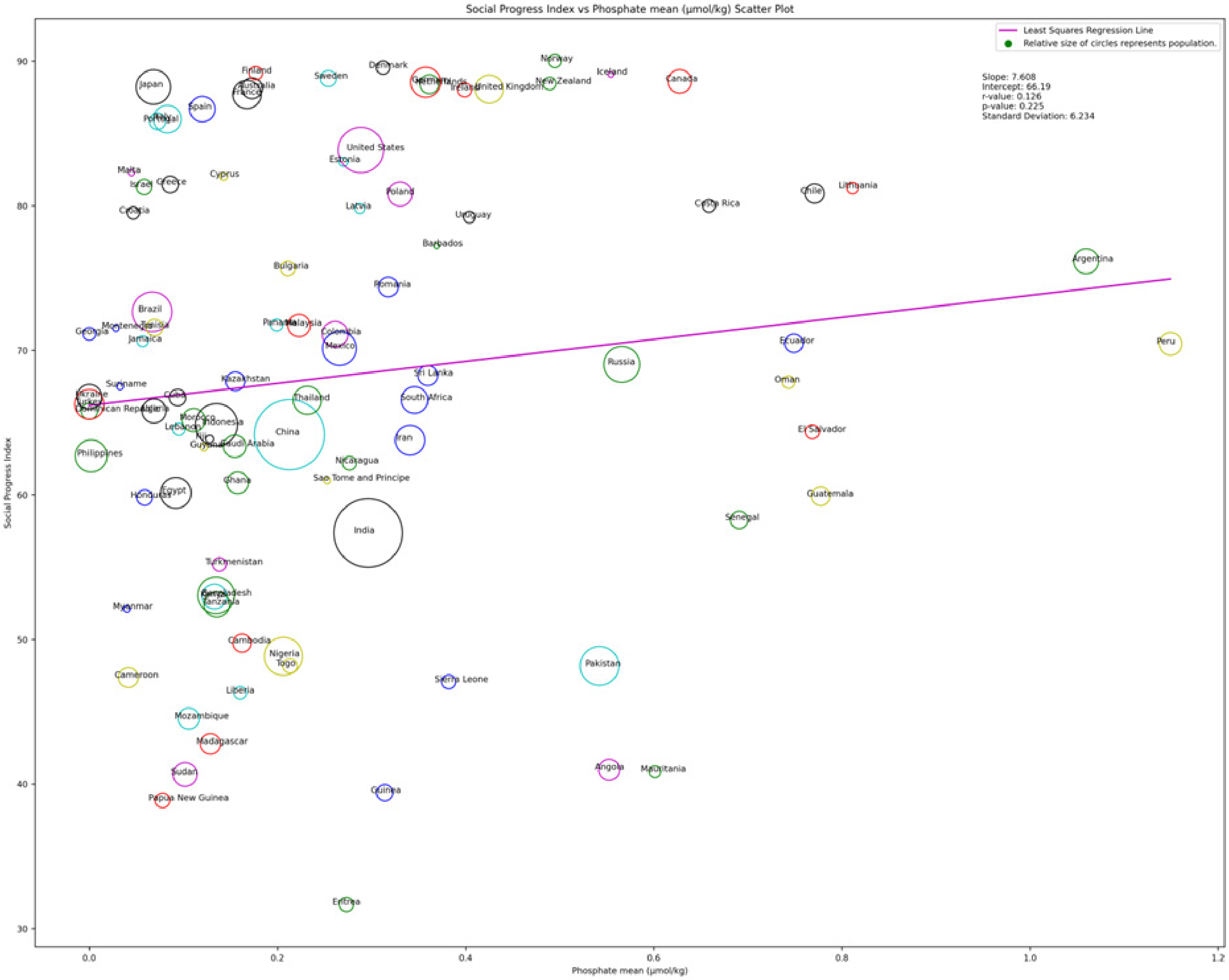


Figure 9. SPI vs. phosphate concentration of all nations shown as a scatter plot.

to basic education.

Limitations

The conclusions that can be drawn from the data are limited by various factors. It was found that areas outlined in the SPI and EEZ data do not correspond exactly, and therefore correlations between the social progress data and oceanic data are inaccurate. For example, American Samoa is identified as a unique area in the EEZ data, but no such distinction exists within the SPI. This may have impacted results, as the environments in these territories may not be similar to the socioeconomic environment in the mainland nation. In addition, it was found that large amounts of phosphate data did not fall into any nation's EEZ at all, which greatly reduced the amount of data available for analysis. Furthermore, some nations possessed few data points, with the Democratic Republic of the Congo, Syria, and Romania all having only one phosphate measurement in their EEZ. Thus, the values used in this analysis may not be representative of the true phosphate values for these nations. Additionally, the methods used in our analysis had some limitations. The fifth degree polynomial model, for example, tended to give anomalous results when used on measurements that were missing many of their data points. This can be explained by the following. Almost all locations show a particular relationship: ocean phosphate concentration has a positive relationship with depth between 0 m and 100 m, but has almost no relationship with depth below 100 m. For these locations, the polynomial model would use a line of best fit similar to that of Figure 2 and correctly predict the 10 m concentration. For locations missing many data points between 0 m and 100 m, the model would use a near horizontal line to be the line of best fit and overestimate the phosphate concentration at the 10 m depth.

The two areas studied were not controlled for nuisance variables, and while the SPI and NOAA are reputable data sources, some variables on the SPI such as quality of electricity supply are not concrete measurements but rather informed ratings. Thus, it is possible that one or more of the correlations studied were influenced by a nuisance variable or a random variable that was not captured by these data sources (e.g., measurement error). These errors can be overcome using richer datasets and more sophisticated data analysis methods to predict phosphate values. We propose a data reduction technique such as principal component analysis or partial least-squares discriminant analysis as possible alternatives to further investigate this topic.

CONCLUSION

The findings reveal there is evidence of a correlation between a country's scores on the SPI and oceanic phosphate concentration within its maritime borders. Particularly, the data suggest a link between the economic and social wellbeing of a nation and maritime phosphate pollution. This study's findings could be supported with improvements such as more accurate geographic correlations and controlling for lurking variables. Future research should examine specific relationships between a measure of wellbeing such as employment rate and maritime pollution. Maritime pollution greatly impacts the UN's Sustainable Development Goals, with the goals of creating sustainable communities and protecting life below water being especially important. Moving forward, developed nations should lead by reducing their pollution outputs if the world is to achieve its Sustainable Development Goals by 2030.

REFERENCES

1. Fertilizers. How many people does synthetic fertilizer feed? Our World in Data; 2019 [cited 19 December 2019]. Available from: https://ourworldindata.org/fertilizer
2. Crop Yield. How have crop yields changed over the long-term [Internet]. Our World in Data; 2019 [cited 19 December 2019]. Available from: https://ourworldindata.org/crop-yields
3. Phosphates in the Environment. What are phosphates [Internet]. Water Research Center; 2019 [cited 19 December 2019]. Available from: https://www.water-research.net/index.php/phosphates
4. Gollin D., Parente S., Rogerson R. (2002). The role of agriculture in development. American Economic Review, 92(2), 160-164. doi:10.1257/000282802320189177
5. Timmer C. P. (1988). The agricultural transformation. Handbook of Development Economics, 1, 275-331. doi:10.1016/s1573-4471(88)01011-3
6. Warmer H., and van Dokkum R. (2002, March). Water pollution control in the Netherlands [PDF]. Lelystad: RIZA Institute for Inland Water Management and Waste Water Treatment.
7. The Social Progress Imperative. Social Progress Index [Internet]. The Social Progress Imperative; 2018 [cited 18 December 2019]. Available from: https://www.socialprogress.org/download
8. About the Sustainable Development Goals [Internet]. United Nations Sustainable Development. 2020 [cited 29 July 2020]. Available from: https://www.un.org/sustainabledevelopment/sustainable-development-goals/
9. National Centers for Environmental Information. Statistical mean of phosphate on 1° grid for all decades [Internet]. National Oceanic and Atmospheric Association; 2019 [cited 18 December 2019]. Available from: https://www.nodc.noaa.gov/cgi-bin/OC5/woa18/woa18oxnu.pl?parameter=p
10. Marine Regions. Maritime Boundaries [Internet]. Marine Regions; 2019 [cited 18 December 2019]. Available from: http://www.marineregions.org/downloads.php
11. Gollin D. Agricultural Productivity and Economic Growth. In: Pingali P, Evenson R. Handbook of Agricultural Economics Volume 4 [book on the Internet] Amsterdam: Elsevier; 2007 [cited 29 July 2020]. p. 3825-62. Available from: https://pseudoerasmus.files.wordpress.com/2017/01/gollin-2010.pdf
12. Reimers M., and Klasen S. Revisiting the Role of Education for Agricultural Productivity (January 2013). American Journal of Agricultural Economics, Vol. 95, Issue 1, pp. 131-152, 2013, Available at SSRN: http://dx.doi.org/10.1093/ajae/aas118
13. Barro R., Lee J-W. A New Data Set of Educational Attainment in the World, 1950-2010. National Bureau of Economic Research. 2010. doi: 10.3386/w15902