May 12, 2020

Richard She
Global Research Institute
College of William & Mary
123 Sesame Street, NY 10128

To whom this may concern,

It is no secret that pollution has been reaching critical levels within these past decades and different ethnic, or socioeconomic groups across the globe have been marginalized and subjected to face the worst of these conditions. While this has been seen within the United States, for example, in Morrisonville where the presence of the Dow Chemical Company has made it 800 times more likely for people in the surrounding communities, composed predominantly of people of color, to get cancer, studies have shown it may be worse in less progressive nations. It is thus imperative that the process of creating an environmental justice map be implemented immediately in countries such as Vietnam where rampant corruption and censorship bar the marginalization of different socioeconomic groups. This is why my team and I are requesting a small grant of $2,500 to speed us up in the process of creating such a map that can identify areas of environmental injustice in even the most rural parts of Vietnam so ensure that no group is being oppressed.

This research project hopes to collaborate with the United Nations Statistical Division, more specifically the United Nations Global Geospatial Management sub-committee, to further develop and implement our environmental justice map after our prototype is finished down the road. We could propose questions like whether or not the Tay ethnic minority is more likely to live near coal plants and industrial factories? Do the emissions and output from these factories lead to higher rates of lung disease or asthma in the area? This is an excellent opportunity to potentially forge a strong path of sustainable development for all nations using information drawn from this map and our team can't be more ecstatic to hopefully change the world.

We are committed to the institute motto that "the secret of change is to focus all your energy not on fighting the old but on building the new" and are driven to bringing environmental justice mapping into the forefront of the community and the world in a new direction.

Sincerely,

Richard She

Richard She

Preliminary Mapping of Environmental Justice Through the use of Geospatial and Correlation Methods

---

Problem Statement

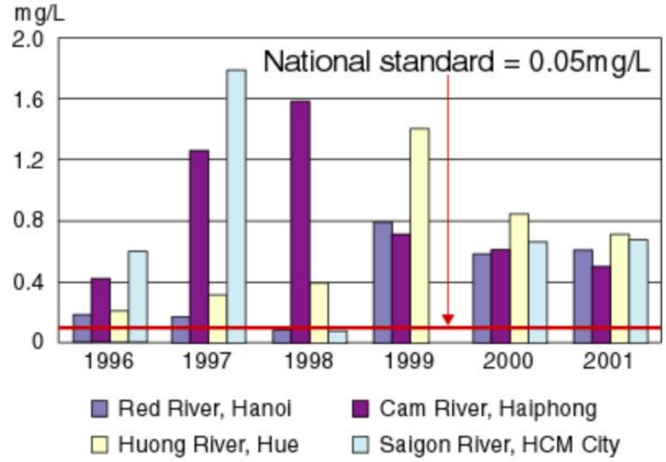
Introduction:

What is Environmental Justice? Coined in the 1980s Environmental Justice emerged as a prominent concept in the US as a social movement that focuses on the fair distribution of environmental benefits and burdens. Its fundamental goal is the fair treatment and meaningful involvement of people regardless of social factors including but not limited to race, or income in the development and implementation of environmental laws. With environmental degradation and inequality on the rise, the idea of Environmental Justice has continued to be pushed into the forefront of the world. Several decades, it wasn't uncommon to find that a disproportionate number of polluting industries, power plants, or waste disposal areas were located near low-income or minority communities. While this proportion has significantly reduced in many developed countries it remains a problem for many developing nations. The general umbrella of environmental injustice covers many aspects of community life and can involve burdens that compromise the health of residents. For example, inadequate access to affordable and healthy foods can even be considered under this umbrella, meaning environmental justice spans a wide and elaborate set of definitions.

Many prevalent examples of environmental injustice exist today. A prime example of this is the water crisis that took place in 2014 in flint Michigan where the city switched its water supply from Detroit's system to the Flint River in a way to conserve money. They failed to consider that for over a century the Flint River was treated as a waste dumping ground for local industries, from car factories to meatpacking plants. The Flint River has even caught fire twice. In Morrisonville, US, the presence of the Dow Chemical Company has made it 800 times more likely for people in the surrounding communities, composed predominantly of people of color, to get cancer. Many other instances of long term injustice take place daily around the globe however they aren't all as well documented or apparent in many developing nations. For example, in 2016, an extensive case of marine pollution left immense damage to the fishing industry and the surrounding community in Vietnam. In only four provinces alone, an estimated 300 tons of fish died, leading to the transnational lawsuit of nearly 8,000 citizens. Environmental Justice truly is an extensive issue due to the fact that inequality is prevalent globally. Inequality exists everywhere, with it being more prevalent in certain areas than others and certain governments addressing it more than others as well. Why is this environmental inequality such a big deal many may ask? Well, its a breeding ground for corruption and it oppresses the freedom that human development is supposed to attract as defined by Amartya Sens. Human development cannot be measured by simply the advancement of the middle or upper class but rather than an entire system. When we can create a system where all people are having their choices and freedoms improved then we are truly digging into the notion of human development. It is thus imperative that we address inequality where every it may stand in order to not only develop infrastructure and institutions sustainably but also to expand these "freedoms" to all persons.

Looking at how these issues affect Vietnam solely, the decades of rapid industrialization has come at a high cost of the environment with Vietnam now being ranked among the top 20 countries with the worst air quality. Furthermore, Vietnam is located downstream of major rivers and thus contains many vast river networks, however, the lack of civic awareness and industrial waste management programs have led to a large deal of water pollution where the people's main source of water comes from rivers and streams. As a result, in many places, especially remote locations, diseases such as e coli and cholera run rampant [1]. In a study case in 2009, researchers found extremely high levels of e coli in groundwater which was attributed to the bird flu outbreak from 2004 as people were improperly disposing of poultry by throwing them into rivers and burying them underground [1]. While the government intends to improve water quality in urban areas significantly by 2025, there is no reliable methodology or source to even monitor this goal. The example I gave provides a view on a small facet of environmental injustice that doesn't even bear to scratch the surface of the gravity of the issue. When we factor in other examples of water pollution or even look at other major agents of environmental degradation such as air pollution, which is attributed to the deaths of an estimate 60,000 people a year in Vietnam, or deforestation and analyze that along with the social demographic agents of poverty and education, the bigger picture of the injustice can be seen. It is fundamental to note that the actions of the government have failed to mirror what is desired of the people; "following decades of scarcity under the command economy [2]" of Vietnam a significant portion of the Vietnamese people actually began supporting trade-offs of economic development for greater environmental protections with public awareness about the issue of industrialization greatly growing in prominence today. In a survey conducted in 2016, sentiments continue to mirror that of the past where 77% of respondents disagreed with the Vietnam government's approach to "grow now, clean later [2]." Many even recognize the lack of environmental justice studies being conducted in Vietnam, leading to the rise of environmental activism and protests across the developing country. The roots of environmental activism first took its place in the 60s and 70s when concerns over Agent Orange, the herbicide used in the Vietnam War, arose. Since then numerous disputes have erupted over the extensive air and water pollution being produced from industrial processes and various developmental projects. While the statistics behind environmental protests are not available, like much other information in Vietnam, studies by the UNDP Justice Index illustrate that environmental concerns make up a common function of administration complaints [2].

Emission standards continue to be violated, evident of this figure below and factories continue to expand evident of the background surrounding the Song Lam Sugar Company whos smokes has easily turned "a wet white shirt into a dry grey one [3]" from the reports of the villagers, most of whom are of lower-income, surrounding the area. It's important not to forget to mention that the officials running such a company have refused to negotiate with the protesters of the company. This is but one of many issues that go unaddressed in Vietnam and my project hopes to combat such issues. If we identify such areas of inequality and injustice and prove that there is a significant correlation between the pollution from such areas and the disproportionate adverse effects have on certain groups. From there, proper action can be taken to resolve the issue. This image above is a visual regarding the NH4 levels found in the major rivers of Vietnam in the late/early 20th and 21st centuries where the findings are contributed through the dumping of an estimated 93 tons of waste daily [4].



Source: *NEA, SOE reports 1997-2002.*

The human developmental process here really draws itself not only from the environmental situation of the country but also its political side as well. Since we are essentially trying to find and analyze information factored from both environmental indicators such as air quality or water quality and socioeconomic indicators such as ethnic density and income many other institutions can play a role in such a project. For example, the location of industrial factories and polluting plants can allow us to identify where these environmental indicators stem from in respect to the location of these socioeconomic indicators. While Vietnam suffers from a lack of information transparency, luckily a new census conducted in 2019 is being finalized to be released hopefully by the end of the second or third quarter of 2020 with much information needed to supplement such socioeconomic indicators. Other information such as the environmental indicators can be drawn from the air quality monitoring sites across the cities in Vietnam.

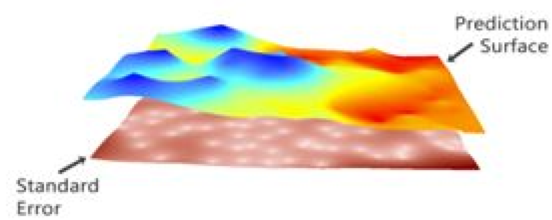
The basis for this project fits under the descriptive inquiry as it seeks to create a profile of a set of extensive social and environmental indicators within a nation, drawing from those indicators to find correlations between different variables that exist. The information used can fit under many other types of inquiries as well with the correlations coming out of this project capable of being used as an explanatory inquiry, in giving hypotheses on why marginalized groups face the conditions they face and other information of such nature. Beyond this, we are trying to introduce the use creation and use of an EJ map in places such as Vietnam to ultimately aid in their sustainable development covered by SDG goals 10 which strives to reduce inequality and 13, climate action to ensure that no one group is marginalized to face deadlier health risks or conditions from pollution. We could propose questions like whether or not the Tay ethnic minority is more likely to live near coal plants and industrial factories? Do the emissions and output from these factories lead to higher rates of lung disease or asthma in the area?

Research on how to translate this idea into a sophisticated visual map has sparked since the emergence of the 21st century in a few developed nations, however, it has failed to find its ground in developing countries where such information is vital for sustainable development planning. Prototypes of these EJ maps have been created in the Netherlands and the US but from what I know it still doesn't exist in Vietnam. This might be partially due to the lack of concrete data in Vietnam preventing many of these environmental and social indicators from being collected. The goal of this project is the address the various factors required to create an extensive Environmental Justice map and its role in sustainable development. As such, the information covered in this paper will explore a combination of methods from different paths in aggregating and correlating various data points within developing nations such as Vietnam to identify potential instances of injustice taking place. A heavy focus is presented to the role of Kriging & population estimation techniques, and various forms of correlation processes such as Moran's 1, an autocorrelation technique, and the Pearson Correlation in studying specific social and environmental indicators. Other methods to address various existing problems in EJ mapping and potential solutions to such issues will be discussed.

Literature Review

Kriging Methods:

Mapping EJ is quite the complex process I chose to focus on presenting methods to aggregate air pollution with the social indicators to find if certain areas are more prone to health risks. When determining the best methods to use to estimate air pollution indicators, we must consider the resources available in a country. So a One Size Fits All approach will not for everyone. For example, Vietnam suffers from a national shortage of air quality monitoring sites due to poor management and lack of maintenance, making it hard to track the flow of Ozone, SO2, or Particulate Matter pollutants. In order to address this issue, I found it imperative to implement the use of Kriging techniques, most notably ordinary Kriging. Kriging is a geospatial interpolation technique that uses a series of complex mathematical formulas closely relating to regression analysis which it uses to estimate values at unknown points or locations based on the values at known points [4]. Ordinary Kriging is a two-step process in which it first creates variograms and covariance functions to estimate statistical dependency which it then makes a prediction for the unknown values of [4]. Variograms are functions that describe the degree of spatial dependence, or the relationship between the



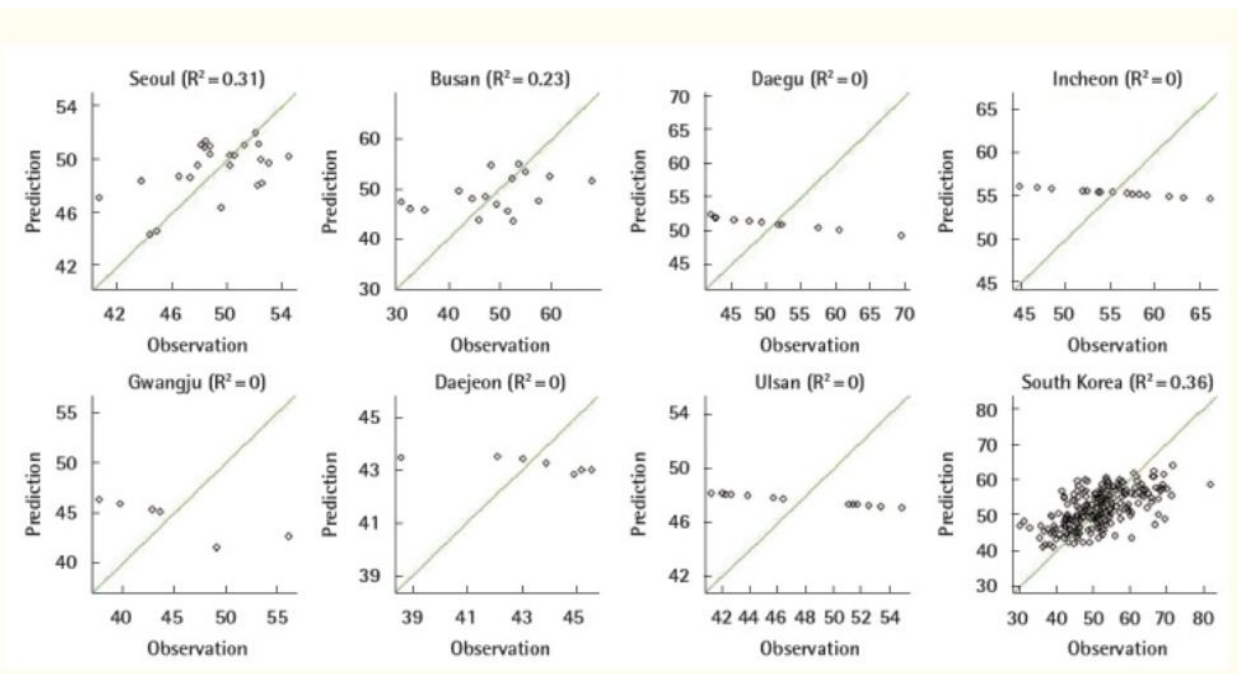
values of variables or locations, of a field. In addition Kriging, otherwise known as Gaussian process regression, and within a normally distributed confidence interval, Kriging gives a least-squares estimation. Since its formal introduction in the 1960s by Georges Matheron, Kriging has since become a core component of geostatistical methods [4]. Kriging was originally developed for mining but has since seen its uses in environmental science and geostatistics. It was originally used to avoid bias in measuring the average grade of mining panels but has since received many progressive applications in the domain of earth sciences, and even in the analysis of computer experiments. Today its applications find itself in the design of aircraft, predictions of movements in nanomaterial, and even the modeling of social systems. Furthermore, Kriging techniques have played a huge role in studies ranging from investigating groundwater quality to monitoring ozone or particulate matter. The great thing about Ordinary Kriging is that when it is applied to air pollution mapping, we can better predict long-term concentrations of Particulate Matter, Ozone, or SO2 values which compile the most significant portion of air pollutions emitted in places such as Vietnam. A significant paper details the use of Kriging techniques in determining the relation between the spatial distribution of sulfur dioxide, ozone, and particulate matter and asthma in New York [5]. By comparing asthma hospitalizations with air pollutants, major threat areas can be identified.

Air quality data and pollution concentrations of SO2, O3, and PM2.5 were collected from close to eighty monitoring sites around New York, and then using Kriging methods, scientists were able to create a spatial distribution map detailing predictions of such concentrations across the entire state. Through the use of Ordinary Kriging, scientists developed a dataset detailing the parameters of a semivariogram including the Mean Standardized and Root Mean Square Standardized error [5].

Items	**Minimum**	**Maximum**	**Mean**	**Standard Deviation**
2005				
Asthma Rate	37.77	94.05	53.04	8.39
Ozone	29.56	44.48	40.21	1.17
$PM_{2.5}$	7.15	16.15	11.39	1.15
SO_2	2.58	18.65	8.46	2.88
2006				
Asthma Rate	37.34	98.30	52.46	9.77
Ozone	30.49	41.92	37.42	1.21
$PM_{2.5}$	5.68	13.30	8.73	1.05
SO_2	2.14	18.69	6.91	2.31
2007				
Asthma Rate	36.56	94.87	50.96	10.43
Ozone	31.75	42.69	39.08	1.76
$PM_{2.5}$	5.73	15.91	9.49	1.08
SO_2	2.81	13.79	7.18	2.38

Asthma rates from discharge hospital reports were compared with extracted values of pollutants on a country by country basis and correlation analysis from the SPSS statistical program was used along with the Pearson two-tailed correlation analysis to provide numeric values of correlation among each pollution concentration (O3, SO2, PM2.5) [5]. The study found significant associations between SO2, PM2.5, and asthma levels. Even though this study was conducted in New York State, the methodology used can greatly aid in Vietnams case. About five percent of the population in Vietnam suffers from asthma with only 40% being able to manage it effectively. If this methodology is further analyzed with income or ethnic distributions in different provinces, we can more effectively tackle environmental injustice; evidence may show that certain groups are more susceptible to such diseases due to their location and the facilities surrounding them.

A similar approach was investigated in South Korea where the PM10 concentration across 226 monitoring stations was used to derive ordinary Kriging models for each of the seven major cities and the nation [6]. The monitoring sites were chosen from a variety of locations not only including places in these urban resident buildings but also places near roads and national background networks. By taking monitoring sites from a diverse dataset made from Ordinary Kriging can become more accurate. For each monitoring site, daily averages of PM10 concentrations were derived from hourly readings to obtain an annual average. In order to compute such annual averages, the researchers included sites for which there was at least one daily average per month for more than 9 months and less than 45 consecutive days of missing daily averages [6]. The predicted annual average of PM10 was calculated along with the R2 and Mean Standard Error which tells us how close the data are to a fitted regression curve so basically how strong the correlations are.



The results obtained from this research project perhaps weren't the best evidence of the R2 values in this graph, however, this provides us important data on the limitations of such techniques and how we can circumvent them. In addition to their use of a simple mean structure, the researchers also noted that the small number of monitoring sites and the presence of insufficient spatial variability would have poorly affected the performance models in use. Within the five cities with an R2 value of 0, there were fewer than 15 monitoring sites and the spatial variation of PM10 was relatively low compared to other cities [6]. It is therefore imperative that we develop exposure models that accurately depict the level of spatial variation in these air pollution concentrations. This seems to fall under a general umbrella that I would like to describe as "garbage in garbage out" meaning that the presence of good data comes first. Regardless of how good our models and methods may be, our results can only be as good as our initial data sets. Without a diverse and generous dataset, our results will be skewed and will not accurately portray the state of which it is like with what happened in this previous paper. Likewise, for the implementation of Kriging in New York, many results were surprisingly inconclusive or negatively correlated due to the fact that asthma data was not readily available or accurate and that there was a lack of uniformly distributed pollution and health data for the researchers to work with. This is something we must be wary about when we apply Kriging Techniques in Vietnam. Effective mapping and recording of such diseases like asthma are extremely limited in Vietnam, however, with little to no data exists in the more rural areas. The results and use of Kriging can only be so specific when we are marred with this issue.

Moran's 1 (Auto-Correlation):

While improper datasets may prevent an effective and precise version of the environmental justice map from coming to fruition in Vietnam, given the existence of proper datasets, the unequal exposure of substances can be analyzed through the use of the local and global Moran's 1 statistical methods and Pearson Correlation methods. Moran's 1 is an autocorrelation technique developed by Patrick Alfred Moran and the technique essentially characterizes the correlation between signals among nearby locations. Research on our environmental justice topic was modeled and executed in Ghent, Belgium by Thomas Verbeek found in ***Unequal residential exposure to air pollution and noise: A geospatial environmental justice analysis for Ghent, Belgium.*** [7] Given a set of features and associated attributes, Moran's 1 can calculate whether the expressed value is clustered, dispersed, or random as it is the correlation coefficient for the relationship between some variables and its surroundings.

The Moran's I statistic for spatial autocorrelation is given as:

$$I = \frac{n}{S_0} \frac{\sum_{i=1}^{n} \sum_{j=1}^{n} w_{i,j} z_i z_j}{\sum_{i=1}^{n} z_i^2} \qquad (1)$$

Figure 1: Illustrates the mathematical formula behind Spatial Autocorrelation. An expected and observed index value is calculated where a computed z-score and p-value is used to determine whether the difference between these two indices are significant or not.

Through the use of Spatial Autocorrelation and Pearson correlation values, the author found that neighborhoods with a lower average income, higher unemployment rates, more minority ethnic groups faced greater rates of air pollution but not noise by combining such indicators with socioeconomic values. The data for air quality was derived from the ATMOSYS annual air quality map from traffic-related pollution, providing rasterized georeferenced data. The quantified noise exposure was drawn from various noise maps from the 2010 EU Environmental Noise Directive while other socioeconomic data were collected from 201 statistical sectors which averaged 1,200 responders each [7]. The significance of the Moran's statistical coefficient values for residential exposure to air pollution was computed through the calculation of the p-value of 999 random permutations. The p-value, in this case, measures how likely the cluster value computed from Moran's 1 is to occur randomly. The possibility of pollution randomly being there is thus taken into account. Concluding a value of .78, the spatial distribution of air pollution illustrates a clustered value [7]. Comparing Moran's 1 with the Local Indicators of Spatial Association (LISA) areas of high air pollution clusters were identified. Moran's 1 value for four covariates: median household income, percentage of people of foreign origin, percentage of rental houses, and the number of house moves per 1,000 inhabitants were also calculated and compared with the cluster distribution of air pollution. Multiple regression techniques were also employed to analyze the correlation between covariates. While the results from this study provide robust and strong conclusions, there were still many limitations nevertheless. These limitations came in the form of effectively modeling and the application of models. For example, the covariates of noise levels and air pollution were modeled separately and the models merely resemble an estimate of the real situation at hand. This may lead to certain factors not included in the models to be neglected from the significance tests and falsely attribute the results. Furthermore, the analysis process was cross-sectional and not longitudinal meaning that many casual relationships are being ignored. Similar to the modeling problem, ignoring causal relationships and merely focusing on the correlated relationships can ignore explanatory factors that may significantly suggest a different result. Therefore the data points can give us the results on inequality but can't tell us how they were made.

Centroid Containment Methods and Its Variations:

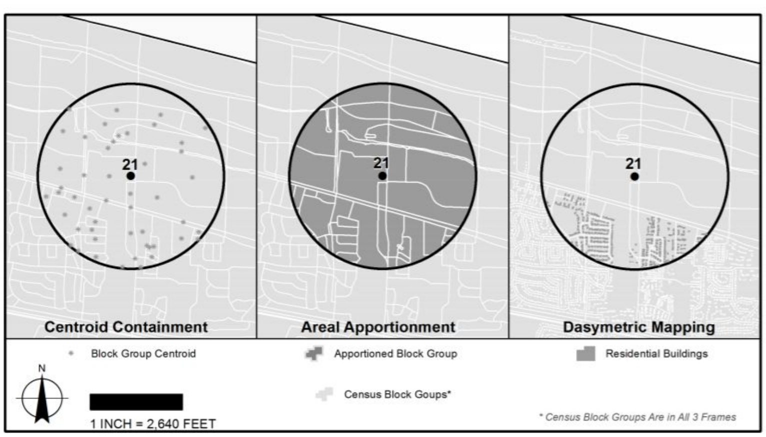
A research paper by Kyle Goodman details the use and comparison of three different GIS population methods for estimating exposure of people to brownfields in Portland, Oregon. Brownfields are former industrial sites that may harbor large amounts of contamination. The author collected three different datasets for this project with demographic information being collected from the US census database in Oregon, contamination zone data being collected from

Oregon's Department of Environmental Quality, and city boundaries being collected from the Oregon Metro. The Oregon Metro contains GIS data from 25 surrounding cities and two counties. The first technique used to estimate the population within the contamination zones is the Centroid Containment Method which, "converts the census enumeration polygons into points within the center of the census block polygon. [8]" The Centroid Containment method is a distance-based method that stems from areal interpolation. The number and characteristics of individuals potentially exposed to a hazard can be estimated by counting points that fall inside a buffer zone. These points are the centroid or geometric center of their respective census blocks. The Centroid Containment Method then creates a buffer zone of a .5-mile radius, which is the circle in the slide, to use to collect the socioeconomic factors surrounding the area of study. The second method is the Areal Apportionment Method which is one of the most common GIS techniques used for population estimation in environmental justice. This method relies on the use of proportional adjustments where the goal is to include the population that is within the exposure zone, excluding those who are not in the zone. An intersect geoprocess is run to capture the census blocks within the contamination zones near the brownfields. To calculate the population within the predicted polygon exposure site the formula in figure 2 is used.

$$Ps = Pt * Pcb$$

Figure 2: Where Pt is the population type is multiplied with the Percentage of Census Block (Pcb) to find the potential people affected by the contamination (Ps)

The third and final method used is the vector-based dasymetric mapping which serves to estimate people per building within the buffer zones surrounding the brownfields. While the other two methods rely on the use of generic shapes, this method provides the use of more realistic spatial units. Dasymetric mapping reorganizes the aggregated data by taking it out of arbitrary geographic boundaries and redistributes it into a more realistic representation of the data (that can then be measured within the affected area.



The number of people per building (PPB) was found by Pt by total building volume (Bt) and multiplying this result by the volume of each individual building (Bv). This gives the formula below.

$$PPB = \left(\frac{Pt}{Bt}\right) * Bv$$

In order to calculate the percentage of minority populations at risk of contamination and ultimately the Population Inequality Quotient Index which is the "proportion of minorities in at-risk areas to their proportions in not-at-risk areas, [8]" the formula below was derived.

$$\left(\frac{At\text{-}risk\ Minorities/Total\ citywide\ at\text{-}risk = At\text{-}risk\ Ratio}{Not\text{-}at\text{-}risk\ Minorities/Total\ not\text{-}at\text{-}risk\ citywide\ population = Not\text{-}at\text{-}risk\ Ratio}\right) = Population\ Inequity\ Quotient\ Index$$

Both the Centroid Containment Method and the Areal Apportionment Method had less than a .03% difference in estimating the population around Portland while the dasymetric mapping had 2.5%, resulting in 9,000 fewer people not being estimated. This was due to the "ancillary buildings data not having exact classification for single-family residential or multifamily residential buildings across the entire dataset. [8]" Overall, however, all three methods had less than 1% difference when comparing the percentage of people affected by contamination. The Areal Apportionment Method tended to overestimate values while the dasymetric mapping method tended to underestimate values.

We can employ these variations of population estimation methods such as the Centroid Containment Method to evaluate the socioeconomic demographics of a population around hazardous location zones such as local factories. Something like this is essential in the making of an environmental justice map since we need to analyze the demographic living around pollution emitting buildings. Such a task was done in Florida using information from the 2000 US census to estimate the race, income, and education in various locations surrounding several local industrial complexes. 71 National Priorities List (NPL) sites identified were used to evaluate the socioeconomic variables within it such as race or income [9]. The researchers received the coordinates of these locations from the EPA and mapped each of these 71 NPL hazardous sites using ArcMap and verified said locations with Google Maps. In situations where the coordinates given by Google Maps and the information given by the EPA were different, the researchers confirmed the centroid location of these sites with the Florida Department of Environmental Protection [9]. The demographic of the area of interest were downloaded from the US Census Bureau and using the Centroid Containment Method and Logistic Regression Analysis, they were able to derive the entire demographic within a 1-mile diameter of these hazardous locations. One of the limitations of spatial studies likes this is that the findings vary with the geography chosen for example using a different census enumeration district would lead to different results as it is the centroid of these districts which are counted within the Centroid Containment Method.

Discussion:

The limitations of each of these studies and of each of these methods are quite the handful to account for when deriving a research plan. From my analysis of the subject, it boils down to having very precise and diverse data that can be modeled well for each differing situation. The Centroid Containment Method relies on the enumeration of census districts and the proper demographic data being quantified while Moran's 1 requires the implementation of models that can best express the real situation of the research area at hand in the proper cross-sectional or longitudinal manner that can account for the existence of both casual and correlation relationships. Kriging also requires the implementation of a wide variety of data points so it can accurately express predictions and output proper analysis.

Research Proposal

Given the correct recourses and opportunities to develop this map can do wonders for not only Vietnam but our international community. Being able to develop such an extensive map for Vietnam can greatly enhance our understanding of the interweavings of the relationship between the environment and socioeconomic factors. Once a basis for environmental justice is established, we can then begin asking why such results exist, and how we can combat it. First of all, are certain socioeconomic groups being marginalized? What kind of adverse health risks do they face? Why are these groups concentrated around these hazardous areas? Are these hazardous areas properly regulated? Why have the complaints of these groups not been addressed? Is it due to under table bribing? Where does this corruption stem from? How can we bring humanitarian aid to these areas of need or how can we address this marginalization? Just thinking about the possibilities this map can begin to bring makes me ever so determined to work on this project with my team. Sustainable development is hard to come by in a world plagued by corruption and environmental issues however my team and I believe that the establishment of a comprehensive environmental justice map can be a true game-changer and integral to future guidance in the development, especially for developing nations such as Vietnam. Bringing awareness of a problem into the light is the first step in addressing it. If we are able to identify marginalization and environmental inequality within Vietnam, we can then bring sustainable development for the people and by the people, ultimately enlarging the freedoms of all people no matter the race or status.

Our plan provides a comprehensive and multifaceted description outlining the course of action our team plans to take in this research process. The facets of the plan are outlined below:

I. Data collection

II. Application of Methods

III. Future Steps

I. Data collection:

The first step to execute in every project of this magnitude is to provide clean and concrete data to which our techniques and methods of applications can return the best results possible. Instead of attempting to apply all of our methods first on a nation-scale, we'd like to first pick a small region of interest to focus on instead, preferably Hanoi. Being the capital of Vietnam, the environment and location of this city provide many favorable qualities to implement our project. Aside from being known for its architecture and delicious food, Hanoi harbors another great distinction: air pollution. In 2016, Hanoi was ranked second among the cities with the worst air quality in all of SE Asia. The surges in construction projects, industrial complexes, and coal powerplants provide ample areas of study to apply our Containment Method. The lack of air quality monitoring sites has also been an issue of constant prevalence, especially in Vietnam, deterring many from being able to properly research the flow and concentration of air pollutants in the developing nation. While the nation has promised to upkeep and collect data from these monitoring sites, the opposite of what was intended has been occurring over the past years. Many of these sites suffer from the lack of maintenance and timely data collection. Luckily Hanoi is getting a second chance with approximately 70 of these stations planned to be installed within this year [10]. This brings the total number of stations to 81 across the city which will greatly aid in the data collection effort to provide clean and concrete data for our project, one of the highest amounts in all of Vietnam [10]. Secondly, Hanoi is a melting pot of both culture and other factors and its diversity that will allow us to obtain demographics with high and visible levels of variation. Having highly variable data will allow us to better distinguish the influence certain groups may have in our final result. Information based on the new and upcoming release of the 2019 census data will provide us the most accurate data possible in Hanoi. With such a high number of stations revamping the air quality monitoring culture in Vietnam, census data and access to other maps and resources rarely mirrored in other cities of Vietnam and a highly diverse population, Hanoi will prove to be one of the best locations to initially base our project off of. Our team anticipates that much of the socioeconomic indicators necessary for conducting this project will be found when the census data is released in the coming quarters. The contribution of 70 monitoring stations will greatly aid in our journey to mapping environmental justice, however, the possible lack of maintenance and data collection of these sites may present a substantial obstacle.

II. Application of Methods:

Once we have our data sets collected in Hanoi, we can begin by analyzing the air quality components surrounding and within Hanoi. Many of the components we will collect come from three categories of pollutants which contribute the most to drops and dips in air quality anywhere else around the world. These components are Sulfur Dioxide (So2), Ozone, (O3) and Particulate Matter 2.5 (PM2.5). SO2 is a gas primarily emitted from the combustion of fossil fuels which is primary not only in Vietnam but across the world. Other things said, cars and such combustion processes also contribute to ozone, PM2.5, and other Volatile Organic Compounds. As such, these three components make up the majority of air pollution globally. Data from these newly established air quality monitoring stations in Hanoi will be recorded hourly to create daily averages over a period of 6 months in order to establish long-term averages for each of these air quality components. The Ordinary Kriging Method will then be applied to create a distribution across Hanoi of the concentration and prediction of each level of concentrations. While this process is taking place, we must ensure that our data is normally distributed, and is stationary. If our data is not normally distributed, we can apply a transformation through the use of log or arcsine until it becomes normal. The normality of our data distribution is essential in the Kriging process. Once we have this strong map of concentrations, we can start to populate it with socioeconomic indicators we have collected through the 2019 Vietnam Census on the demographic surrounding the residence of Hanoi. Hanoi is home to 54 different ethnic groups and a wide income distribution curve. Once we have down, we can begin to find creating our buffer zones which we will set as a 0.5-mile radius circle surrounding industrial factories and complexes. The location of these factories will be aggregated on ArcMap and compared/checked with the actual location of these areas given by a map of such buildings in Hanoi. The demographic concentration of each respective census enumeration block will be found along with its respective centroid. Once the centroid of each block is established, we can collect the demographic concentrations of the census blocks whos blocks' centroid falls within the buffer zone. Finally using Moran's 1 the correlation behind each of these environmental and socioeconomic layers can be found. R2 and p values will be some of the information outputted to which we can then find significance in our null hypotheses. Some concerns we have are whether overlaying predictions derived from Kriging will function well with autocorrelation techniques and whether the accuracy of such data is strong. Concerns on scaling the size of this project from Hanoi to the entire nation of Vietnam are legitimate in the data collection process. The same process will be applied on the national scale however data collection may become an issue. We plan to allocate approximately $40,000 in our initial prototype in Hanoi to set a foundation for future work. Costs will vary depending on how to smooth data collection can go around the rest of Vietnam. The team considers raising money to fund more air quality monitoring sites in not only other major cities but also in the rural areas of Vietnam such as Hoi An or Mui Ne. The worst-case costs for the entire project are projected to be $100,000 which may require additional funds.

III. Future Steps:

Based on some of our previous concerns in proper data collection addressed in sections I and II, we encourage international collaboration on this effort through the creation of an international summit of scientists to explore discuss and improve the models we obtained from our prototype in Hanoi. Here, methods to explore better and even more efficient ways of data collection can be found. It is imperative that cross border collaboration takes place as the issue of inequality and pollution is also a cross border issue all nations are plagued with. Highly correlated information obtained from our prototype is encouraged to be continued to be explored upon by our Vietnamese counterparts in hopes of combatting such issues.

Discussion:

Through the exposure of our work to the international community, we can continue to offer and encourage change and adaption to models that may or may not have worked in our study of Hanoi. Because the credibility of data collection is so skewed in more rural areas we hope that the exposure of our work can incite ways to translate this onto a national scale. As discussed, we hope this information becomes a basis for future research on identifying and addressing inequality in Vietnam. Only through this can we truly confirm the reality of things in Vietnam. If instances of inequality are exposed, we can begin the process of reforming its cause whether it be due to corruption on the local or national level. Humanitarian aid can be accessible through Non-Governmental Organizations such as the Vietnam Red Cross Society, or World Health Organization Vietnam. Only by exposing the problem first can we begin the take steps in combating such issues. While the proposals of many other applicants are impressive, this project introduces a revolutionary side of geospatial mapping that has been rarely discussed or researched compared to other hot topics. The capability of this map represents the entire notion of sustainable development as the information gained from the project can be used as the basis for further research as it has proved time and time again for environmental justice maps. The possibilities for a map of this magnitude are truly endless and can benefit the international community to a whole other degree. While this project may prove to be time-consuming than many others, its benefits in transforming human development greatly outweighs its costs.

Nevertheless, it is an honor to be considered among such a wonderful set of applicants and our team can't be happier to participate in the process.

References:

[1] Person. (2020, February 11). Vietnam culls 30,000 poultry to stem bird flu outbreak - VnExpress International. Retrieved from https://e.vnexpress.net/news/news/vietnam-culls-30-000-poultry-to-stem-bird-flu-outbreak-4053456.html

[2] Gillespie, John, Nguyen, Canh, Nguyen, & Hung. (2019, August 1). From "Weak" to "Strong" Sustainability: Protesting for Environmental Justice in Vietnam. Retrieved from https://online.ucpress.edu/jvs/article/14/3/1/46107/From-Weak-to-Strong-Sustainability-Protesting-for

[3] (n.d.). Retrieved from http://www.wepa-db.net/policies/state/vietnam/surface.htm

[4] Chilès, J.-P., & Desassis, N. (1970, January 1). Fifty Years of Kriging. Retrieved from https://link.springer.com/chapter/10.1007/978-3-319-78999-6_29

[5] Gorai, A. K., Tuluri, F., & Tchounwou, P. B. (n.d.). A Gis Based Approach for Assessing the Association between Air Pollution and Asthma in New York State, Usa . International Journal of Environmental Research and Public Health ISSN 1660-4601

[6] Kim, S.-Y., Yi, S.-J., Eum, Y. S., Choi, H.-J., Shin, H., Ryou, H. G., & Kim, H. (2014, September 22). Ordinary kriging approach to predicting long-term particulate matter concentrations in seven major Korean cities. Retrieved from https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4178540/

[7] Verbeek, T. (2018, December 13). Unequal residential exposure to air pollution and noise: A geospatial environmental justice analysis for Ghent, Belgium. Retrieved from https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6304432/

[8] Goodman, Kyle, "Environmental Justice and GIS: A Comparison of Three GIS Methods for Estimating Vulnerable Population Exposed to Brownfield Pollution in Portland, Oregon" (2017). Geography Masters Research Papers. 19. https://pdxscholar.library.pdx.edu/geog_masterpapers/19

[9] Kearney, G., & Kiros, G.-E. (2009, June 17). A spatial evaluation of socio demographics surrounding National Priorities List sites in Florida using a distance-based approach. Retrieved from https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2708146/

[10] Person. (2020, February 12). Hanoi to set up 70 more air monitoring stations - VnExpress International. Retrieved from https://e.vnexpress.net/news/news/hanoi-to-set-up-70-more-air-monitoring-stations-4051397.html