Examining the Local Effects of Bus Depots

Submitted as partial fulfilment of a Master's of Science in Geospatial Technologies

Brandon D. Couch[1]

[1]University of Washington - Tacoma, 1900 Commerce St, Tacoma, WA, 98402
Email: brandoncouch8@gmail.com

Abstract

Public transit has long played an important role in getting people from one place to another. Under ideal circumstances, public transportation can carry more people efficiently than private cars, reducing traffic and pollution overall. Yet, it does potentially bring up its own issues, such as increased amounts of pollution, noise, and traffic around major transit centers and depots. These local problems negatively affect those living nearby by exposing them to more pollution, which can degrade their quality of life. Because of this, the construction of new transit depots is often met with stiff resistance from residents, regardless of the overall benefits of increased transit. This brings up a dilemma: Cities will need more public transit as they grow, which will require more depots to be built, but the local downsides caused by depots and hubs make it difficult to find ideal spots to put them. This project seeks to investigate the drawbacks of various forms of public transit by analyzing geographic data on air pollution around transit depots and study their negative impacts on nearby residents. Geographically weighted regression was used to compare different types of emissions to their proximity to the nearest bus depots in New York City and to land value. The end result of this project was to produce a model based on existing data to predict the local effects of bus depots in New York City and determine if there is a relation between bus depots, air pollution, and land value.

1. Introduction

Bus depots are an important part of the transit system as they serve as facilities for maintaining and storing a city's fleet of buses. This helps to keep buses working efficiently, which ideally reduces the strain on traffic by taking cars off the streets, which generally helps to reduce pollution due to there being fewer vehicles. However, the local consequences of buses, especially around depots, are not as well known. Most buses, like other motor vehicles, produce emissions that can be harmful to people's health; having many buses in a single area can produce a lot of these emissions which can exacerbate these problems. This can reduce the quality of life of those living near such areas, which can also bring down land value as those areas become undesirable to live in. However, it is not known how significant these impacts are, or if there is any relation at all.

In this project, I sought to examine data on various types of air pollution and relate them to proximity to the nearest local bus depot in order to determine if bus depots had a significant impact on local air quality. I then sought to analyze data on property values, as a general metric for quality of life, and compare them to the aforementioned data on air pollution and bus depot locations, with the goal of finding a relation between bus depots, air pollution, and land value. For my purposes, I used data from New York City to conduct my analysis and used geographically weighted regression (GWR) to compare the variables, determine if there were significant relations, and ultimately generate a predictive model of land value based on local air pollution and distance to the nearest bus depots.

The purpose behind this project is to determine if bus depots negatively impact the air quality of their surrounding area, and by extension quality of life. Regardless of whether there is a significant relation or not, the results could potentially assist transit companies with selection and planning for new depot sites with network expansion. If a correlation between bus depots and pollution is found, then that will guide the transit companies guide where they build new depots, as well as draw attention to the flaws of existing facilities so that they My work to mitigate them. If a correlation is not found, then that opens up more options for bus depot placement, as they may not negatively impact land values after all; it might also indicate that bus depots have a minimal impact on local pollution compared to other sources, diminishing some environmental concerns.

My initial hypothesis at the beginning of this project was that bus depots would have a significant impact on local air quality, and in turn would negatively impact land values on the surrounding area due to reduced quality of life. I assumed this would be the case as I believed a higher concentration of larger vehicles in a specific locale would highly impact the overall air quality in the surrounding environment.

2. Literature Review

As Hovarth pointed out in "Machine Space" (1974), we sacrifice much of our physical space for machines such as motor vehicles, in the form of roads and garages and other spaces where they take priority over people. We also give up some of our comfort as well, by relinquishing our cities to noxious exhaust and traffic noise. Buses, being motor vehicles themselves, are not exempt from this, even if they do reduce the total number of vehicles on the road. In addition to roads and bus lanes, we also give up sizeable amounts of space to bus depots, which function as garages for storing buses that are not in use as well as maintenance centers for repairing and inspecting buses. (Pawlicki *et al.* 2012) The major factors affected by the presence of bus depots that will be studied include air pollution, noise pollution, and their effects on surrounding land value.

2.1 Pollution

Most buses today currently rely on combustion of fossil fuels such as diesel or natural gas to run, just like most other road vehicles. A study by Nylund *et al.* (2004) shows that the main chemicals found in engine exhaust from buses include carbon dioxide (CO_2), nitrogen oxides (NO_x), and particulate matter such as soot and aerosols (PM). These pollutants are more likely to cause lung problems, such as asthma and wheeze, in children under five years old, as said by Patel *et al.* (2011). This study showed that risk increases the closer one lives to stationary sources of pollution, highways, dense intersections, or other sources of heavy traffic. Ozone has also been shown result in increased hospitalizations during the warmer months for severe asthma, especially among children ages six to eighteen. (Silverman and Ito, 2010) Additionally, CO_2 is a greenhouse gas which heavily contributes to climate change, while NO_x and PM contribute to smog and are also irritants that can be harmful to people's health. (Kheirbek *et al.* 2016)

This same study observed that buses running on compressed natural gas can produce lower emissions overall than diesel if they operate with stoichiometric combustion and proper catalysts but can actually produce more CO_2 and NO_x if it is running without either. Overall, a single bus will produce fewer emissions than the equivalent number of cars, due to their higher capacity which allows them to replace multiple cars. (Hodges, 2010) However, some studies (Rivers *et al.* 2017) have suggested that merely increasing public transit capacity without significant overhaul to system may not actually improve air quality, as buses operating at a lower capacity may not actually increase efficiency compared to smaller cars.

One possible way we could measure emissions data is by using traffic data to produce an estimate of emissions (Samaranayake *et al.* 2014) although said model would have to be modified to only use bus traffic data and emission levels. Such a model would be able to map out emissions in real time, though that may not be necessary for this project. Pollution levels can also be measured directly using special equipment carried around in vehicles driving around areas we want to measure. Apte *et al.* (2017) were able to measure the pollution levels around different parts of Oakland, CA by loading air quality sensors onto Google Street View cars, giving accurate air quality data as they drove around over the course of a year. This study was able to reveal relatively stable hotspots of lower air quality, such as along busy streets. A similar study was conducted in New York City by Herndon *et al.* (2005) using air monitors on vehicles that followed buses. These monitors measured found that CNG buses emitted more methane and formaldehyde than diesel buses on average, especially when the buses backfired when quacking changing speeds. Another study by Shorter *et al.* (2005) measured nitric oxide and nitrogen dioxide emissions and found that transit buses do usually emit more nitric oxide emissions than other vehicles, though hybrid buses, buses equipped with continuously regenerating technology (CRT), and buses running on compressed natural gas emit fewer emissions.

Analysis conducted by Kheirbek *et al.* (2014) suggests that larger amounts of air pollution, specifically nitric oxide, nitrogen dioxide, and black carbon are correlated larger amounts of noise pollution as well, due to heavier traffic. Noise pollution is harmful to surrounding communities, as excessive noise can result in various health problems, including hearing loss, sleep disruption, and even cardiovascular disease because of stress. (Basner et al., 2014) While I was originally going to analyse noise pollution as well, I was unfortunately unable to find enough data on it online.

Not all pollution comes from traffic, much less buses exclusively, however. Other sources unrelated to transportation are also responsible for pollution. Research by Lall and Thurston (2006) suggests that half of particulate matter emissions and aerosols are transported in from sources outside the city, carried by the wind. Ozone concentrations can also vary based on wind patterns; a study by Civerolo *et al.* (2007) predicted that increased urbanization of New York would lead to episodes of higher overall ozone but may cause some areas where ozone levels decrease.

This information is important as emissions are one of the major problems from buses that should be measured around areas more prone to heavy bus traffic, such as depots, to evaluate their impacts on the surrounding area. By studying the impacts of air pollution on the surrounding area, it will become easier judge how much of a problem it is around depots and address the problem with potential solutions. Furthermore, a study by Fricker and Hengartner (2001) showed that certain minorities may happen to live closer to more environmentally undesirable sites, bus depots being one of them, then other people, although the results were not conclusive. This analysis will possibly make it easier to judge if environmental discrimination against certain people is a problem and allow us to propose solutions to improve equity.

2.2 Land Value

Air pollution can be detrimental to people's health, which can make places around bus depots undesirable to live in. This can be reflected in the property values of the surrounding area: as poorer quality of life makes people unhappier, said places become more undesirable and property values become lower. (Welsch, 2006) On the other hand, easy access to public transit services can make a place more desirable and thus more valuable, since it will be easier to get to different places, though there may not necessarily be a lot of stops next to depots. A study in Brisbane, Australia by Mulley *et al.* (2016) showed that property values were higher closer to transport infrastructure. Likewise, a study in Hong Kong by Hui *et al.* (2007) showed that property values aren't necessarily impacted by certain environmental factors, and in some cases noisier property was more valuable, although the article points out that preferences of people in Hong Kong are likely different to those in the West. Both studies model transit, environmental factors, and land value using GIS. Models have been developed to relate environmental factors to property values as way to judge the economic impacts of factors that usually do not have a market value, with the most common one being the hedonic model. (Palmquist, 2005) Census data could easily be used to find land value, which could be compared to air and noise pollution as well as proximity to bus depots.

Analysis of vacant land transactions in New York City by Haughwout *et al.* (2008) showed that property values have been steadily rising over time, and that prices tended to be higher the closer a property is to the Empire State Building in the core of Manhattan. This study also points out that environmental problems can also negatively impact a property's value.

Land value is an easy way to judge how desirable property is based on various environmental factors, as property will become more or less desirable based on how happy people are, which is tied to factors that would reduce the quality of life such as air and noise pollution. In this way, it is used as a measure for how heavily an area may be made undesirable by the effects of pollution from nearby bus depots.

2.3 Analysis and Solutions

This project uses geographically weighted regression (GWR) to analyse the potential relations between different types of air pollution and land value based on distance to bus depots. Geographically weighted regression is used because it is more adequate for mapping out data that may have positive and negative relations in different areas than standard linear regression (Mennis, 2006), which could potentially happen with this data. Robinson *et al.* (2013) found that GWR was more accurate for mapping out nitric oxide and nitrogen dioxide than global regression at local levels throughout the United Kingdom. Similarly, Zhai *et al.* (2018) were able to model the concentration of fine particulate matter in Beijing, China using an enhanced version of GWR, which proved useful in a situation where sensors were becoming more sparse, as well as the ability analyse multiple explanatory variables. GWR has also been used examine proximity to transit lines and their effect on land value, as shown by a study Zhang *et al.* (2020) performed in Hartford, Connecticut. Interestingly, in this study transit was shown to have positive effects in some areas, and negative effects in others.

To mitigate some of the issues surrounding bus depots, we can also potentially use a model for evaluating specific aspects of the transit system, such as the one developed by Liu *et al.* (2020) and apply it to factors like air pollution. Air pollution can also be modeled based on traffic levels along street canyons using an urban background model. (Jensen *et al.* 2009). We can then evaluate the efficiency of the system (Li *et al.* 2020) in order find ways to make transit more efficient and greener. Efficiency can be improved by optimizing certain routes and schedules, potentially by arranging them in a grid like pattern (Pemberton, 2020) to make travel faster and decrease the number of buses needed. The location of depots could also be optimized to find more efficient spots that also will not negatively impact surrounding residences, such as by being further away from residential areas, and to improve equity.

3. Intervention Description

The results of this study are published in a series of maps uploaded to the ArcGIS website as a story map. These maps will include the original data sets for pollution and land value compared with predicted models generated through geographically weighted regression. This will allow users to visually compare the model results with the observed data. Additional maps displaying the results of the regression analysis will also map out the standardized residuals for each model so that users can see where the predicted results significantly diverged from the observed data. Darker values will show higher levels of pollution and land value, or in the case of standardized residuals, greater divergence with different hues for positive and negative residuals. Short explanations will be provided for each comparison, explaining how the predicted and observed values matched up and diverged, as well as how they related to the location of bus depots. Plots generated as a result of the analysis in ArcGIS showing relations between variables will also be included in order to further show how certain variables may or may not impact each other. A final raster data set modelling land value based on the observed data for all types of pollution as well as proximity to bus depots will be included at the end, along with an explanation of the accuracy of this model. The raster model will match the resolution and scale of the initial air pollution rasters. The overall formatting of the story map will follow the composition of the "Discussion" section of this paper.

4. Methods

I obtained the raster data for air pollution from the New York City Open Data Portal website. These raster data sets were produced by the New York City Community Air Survey (NYCCAS). This data was modeled based on measurements taken from about one hundred different air sensors located in different parts of the city. For my analysis, I chose to study the data for five different types of air pollution recorded and modeled in the year 2019, as this was the most recent year in the collection of rasters. These types of pollution include black carbon (BC), nitric oxide (NO_x), nitrogen dioxide (NO_2), fine particulate matter, and ozone. The raster data collection also included data for sulfur dioxide; however, since this data set was not as up-to-date as the others, I chose not to use it. In order to make the raster data compatible with geographically weighted regression in ArcGIS, I first converted all of the raster data to point data which also made it easier to calculate distance to the nearest bus depot.

The data for land value by census block group was obtained from the American Community Survey as part of the United States Census Bureau. This data set was initially only downloaded as a table; thus, I had to join the table to a shape file of census block groups based on GeoID. The block group shape file was initially obtained for the entirety of New York State and then trimmed down to just the block groups in New York City by clipping it with another shape file for the five boroughs of the city. (Manhattan, Brooklyn, Queens, Staten Island, and The Bronx) Both of these shape files were also obtained as TIGER/Line data (Topologically Integrated Geographic Encoding and Referencing) from the United States Census Bureau. One minor issue with the mean land value data was that it only recorded information for owner-occupied households and as such did not include values for leased/rental properties, nor for non-residential or commercial properties. Thus, not every census block group had complete data most likely there were no owner-occupied household included. I also used ArcGIS to calculate the centroid of each block group polygon for the purposes of calculating distance.

Data for the location of bus depots in New York City was obtained through as a KML file (Keyhole Markup Language) based on data from "Open Street Map." This was imported into ArcGIS with no issues. For all of the pollution point files, as well as the block group centroids, I used the Near function in ArcGIS to calculate the distance to the nearest bus depot for each point. Proximity to bus depots would be used as my independent variable for my analysis.

After collecting my data and processing it for use, I performed geographically weighted regression on each of the five different types of air pollution using distance to the nearest bus depot as an explanatory variable. The main purpose of this was to test whether or not there was any correlation between air pollution levels and the location of bus depots. For this analysis, I used a Continuous (Gaussian) Model Type due to the air pollution values being widely distributed, meaning they covered a broad range of values. I also used a Distance Band neighborhood type, which was calculated via the Golden Search Method in ArcGIS. I used a Gaussian Local Weighting Scheme for assigning the importance of different values with points having less of an influence the further away they were in proximity. The outputs of these analyses showed predicted values of each type of pollution as well as plots showing the relations between distance and each type of pollution and maps for the standardized residuals between predicted values and observed values. This would allow me to see how accurate the predicted results were and, more importantly, determine whether or not there was a strong relation between proximity to bus depots and air pollution.

The final analysis would involve using geographically weighted regression to model land value based on all types of air pollution and proximity to bus depots. However, before I could perform this analysis, I needed to translate the point data for pollution over to the census block group data. I did this by aggregating the pollution points into the census block group polygons. This means that if a point fell within a block group polygon, then the value of that point would be assigned to that polygon; if multiple points fell within a polygon, then the mean value of those points would be assigned. Geographically weighted regression was then performed on this unified census block group file using the same modelling parameters as before. This analysis also generated a point-based model based on a single point data-set produced by spatially joining all of the pollution point data into a single file. This unified point data-set contained all the explanatory variables needed for generating a prediction as ArcGIS can only pull explanatory values from a single file. The predicted data-set was converted back into a raster file of the same resolution as the original pollution rasters in order to allow for easy visual comparison between predicted land value and air pollution. The raster model would also allow me to predict the land value where mean land value was not recorded in the census block group data. As with the previous analysis, scatter plots were generated showing the relationships between different types of variables as well as maps of standard residuals. These would ultimately allow me to determine if there was indeed strong correlation between land value, bus depots, and air pollution.

5. Discussion

5.1 Initial Observations

Looking at the pollution raster models produced by NYCCAS, I can already observe a few patterns in where pollution appears to be concentrated. In Figure 1, black carbon is seen to be concentrated heavily in Manhattan, with another significant concentration around Grand Avenue Depot in Queens; this depot is located near Newtown Creek, an estuary surrounded by industrial area known for being heavily polluted. A smaller hotspot can be seen in Hunts Point north of the New York Post depot, which is also home to a lot of industry.

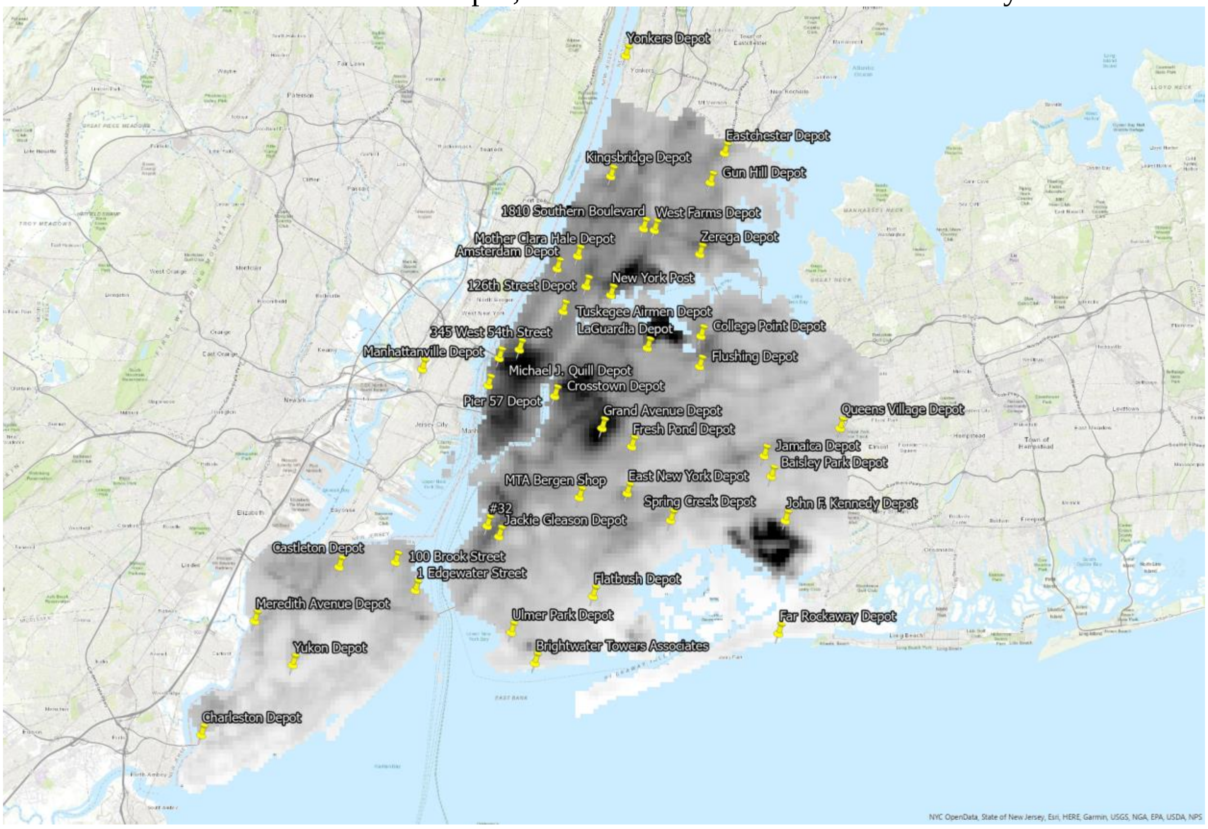

Figure 1: Black Carbon Emissions in New York City, 2019.

More hotspots are observed near the two airports (LaGuardia and John F. Kennedy), each of which also have their own bus depots, although it is likely that these spots of pollution may be more a result of the airports themselves rather than the bus depots near them. One can also see that there aren't very many significant hotspots for black carbon emissions at other depots, especially ones further away from the center of the city in the eastern parts of Queens and the southern portions of Staten Island.

The most significant hotspot for nitric oxide emissions is again located around the Grand Avenue Depot, as seen in Figure 2, matching up with a similar hotspot observed in the black carbon data. However, this hotspot is larger and more intense than the hotspot observed for black carbon. Another fainter hotspot can be observed at Hunts Point near the New York Post depot, although this one isn't as intense as the hotspot for black carbon.

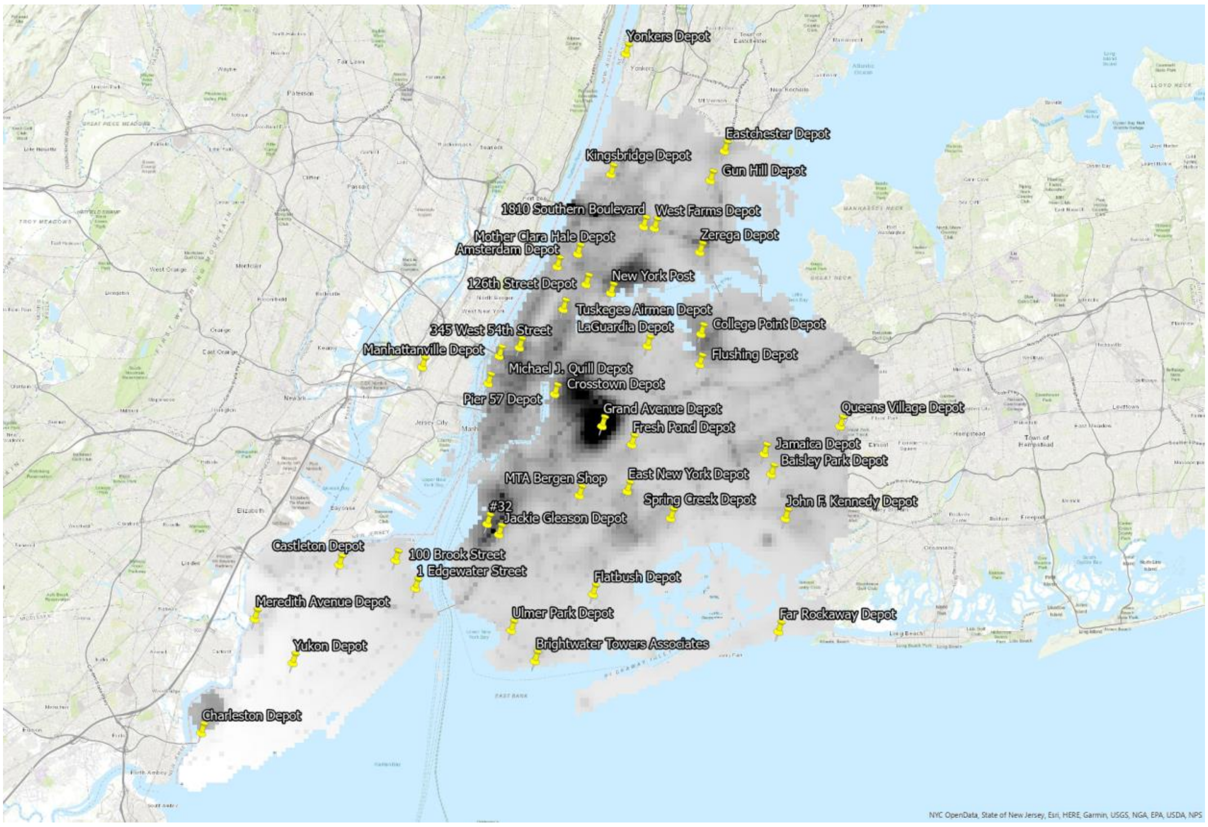

Figure 2: Nitric Oxide Emissions in New York City, 2019.

Outside of those two hotspots, nitric oxide emissions seem to be relatively mild elsewhere in the city, with even lower levels observed in Staten Island, though a mild warm spot can be seen by the Charleston Depot. It is interesting to note that darker lines can be observed in both rasters that correspond to major highways, which would make sense due to the higher amount of traffic.

In Figure 3, nitrogen dioxide levels are shown to be at there highest in central Manhattan, with fairly moderate to high levels throughout the city except in Staten Island, where levels are lower. Hotspots can be seen near Grand Avenue and New York Post depots as in the previous emissions, but do not stand out as much do the heavier amounts of pollution all around. Again, a milder hotspot appears north of the Charleston Depot on Staten Island.

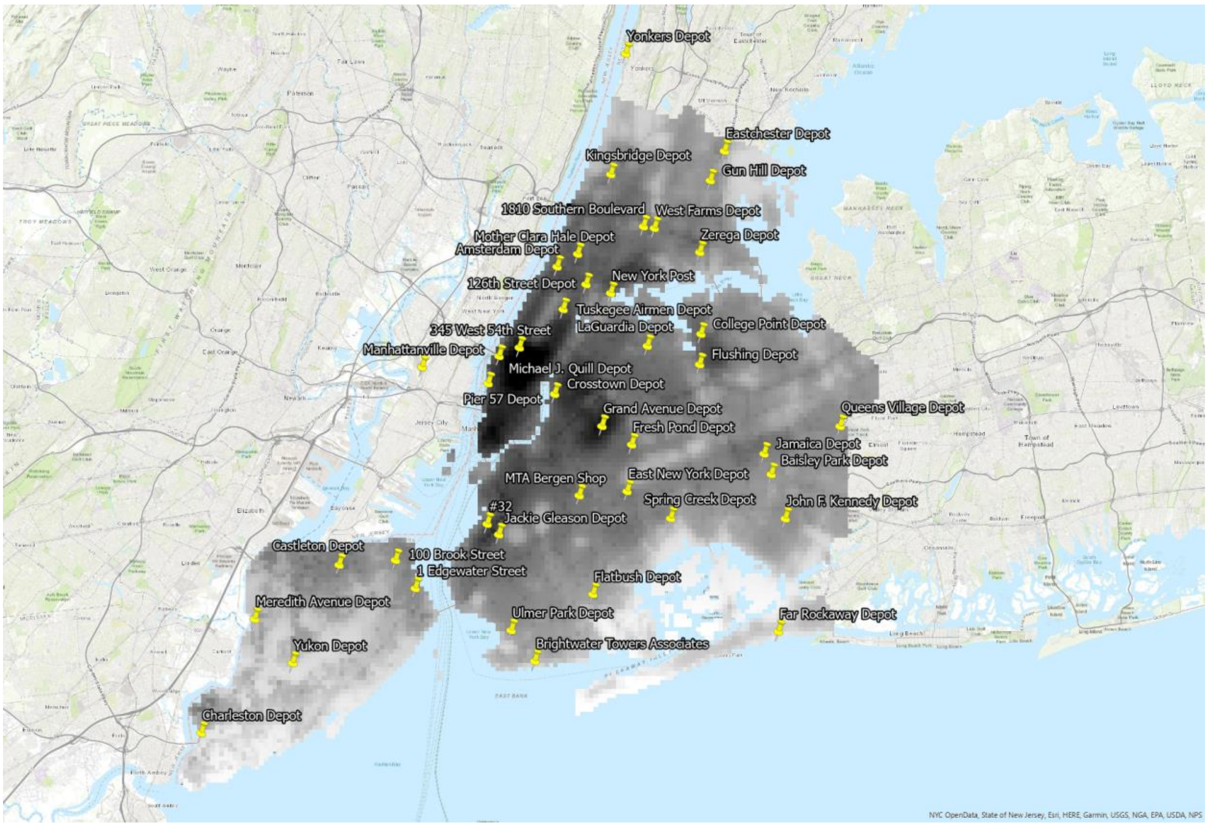

Figure 3: Nitrogen Dioxide Emissions in New York City, 2019.

Like with black carbon, fine particulate matter appears to be mostly concentrated around central Manhattan and Grand Avenue Depot, which can be seen in Figure 4. Levels appear to be lower throughout the rest of the city, with the lowest levels in Staten Island and around Jamaica Bay, with another warm spot north of the Charleston Depot.

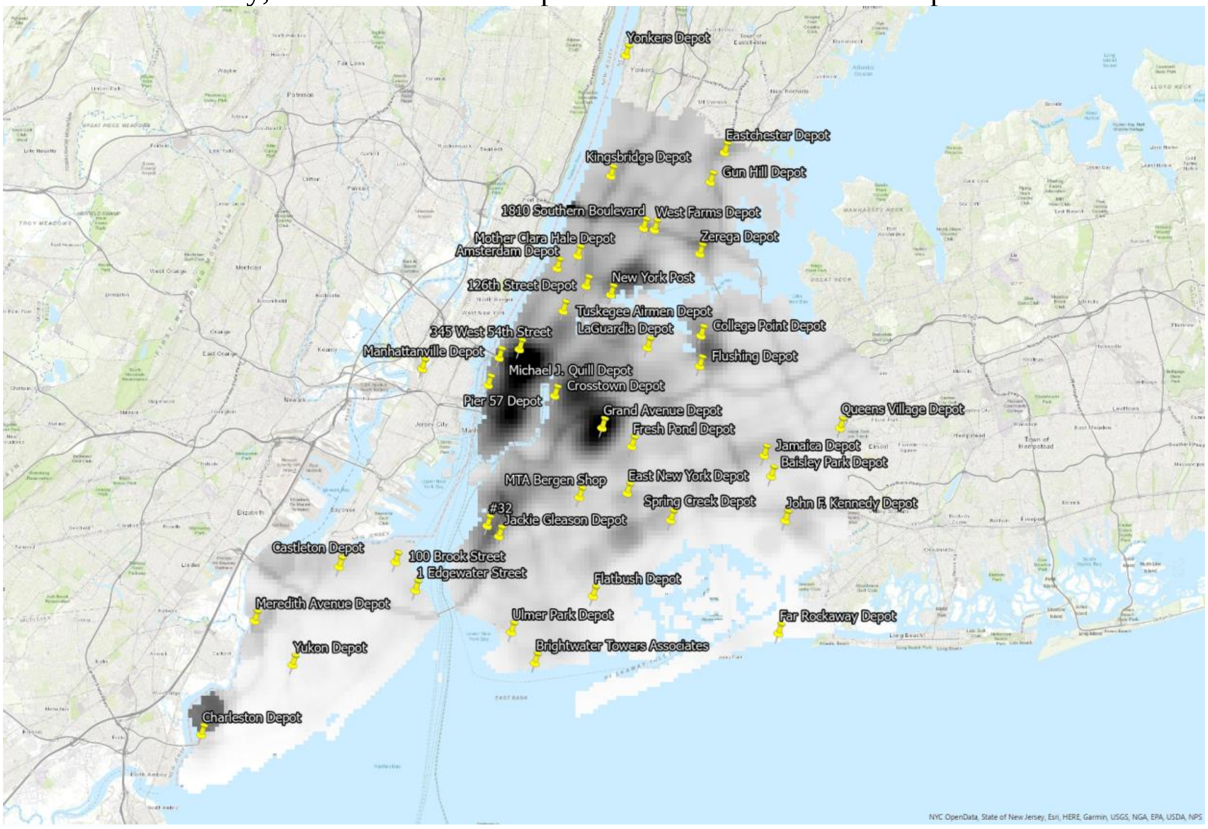

Figure 4: Fine Particulate Matter Emissions in New York City, 2019.

Ozone levels, as seen in Figure 5, oddly seem to be an inverse of other emission levels, with lower levels in central Manhattan and higher levels around Jamaica Bay, which seems to contradict expected patterns. (Higher levels of ozone corresponding to higher levels of other pollution).

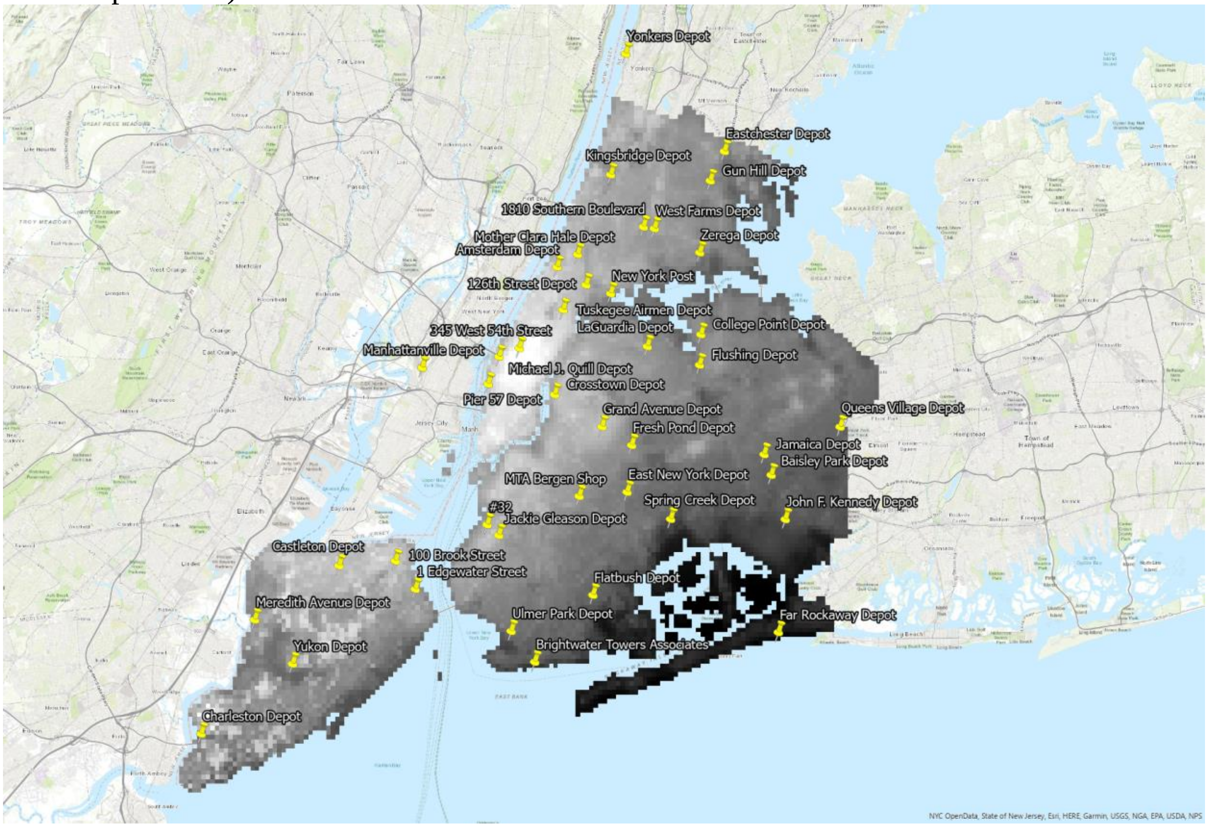

Figure 5: Ozone Emission Levels in New York City, 2019.

My assumptions are that the model predicting ozone levels were off, the data was somehow inverted, or ozone levels for other areas really do have an inverse relation with other pollutants. It is worth noting that the ozone levels are only the average for summer, while the other emissions are year-round averages.

Mean property values appear to be at their highest around Manhattan, with properties becoming less valuable towards the edges of the city, as seen in Figure 6. The least valuable properties appear to be in Staten Island, and The Bronx. Properties also seem to be less valuable around Jamaica Bay. It looks like homes are valuable where there appear to be a higher concentration of depots; however, this may be due to the fact that it is closer to the center of the city where there is a higher concentration of population.

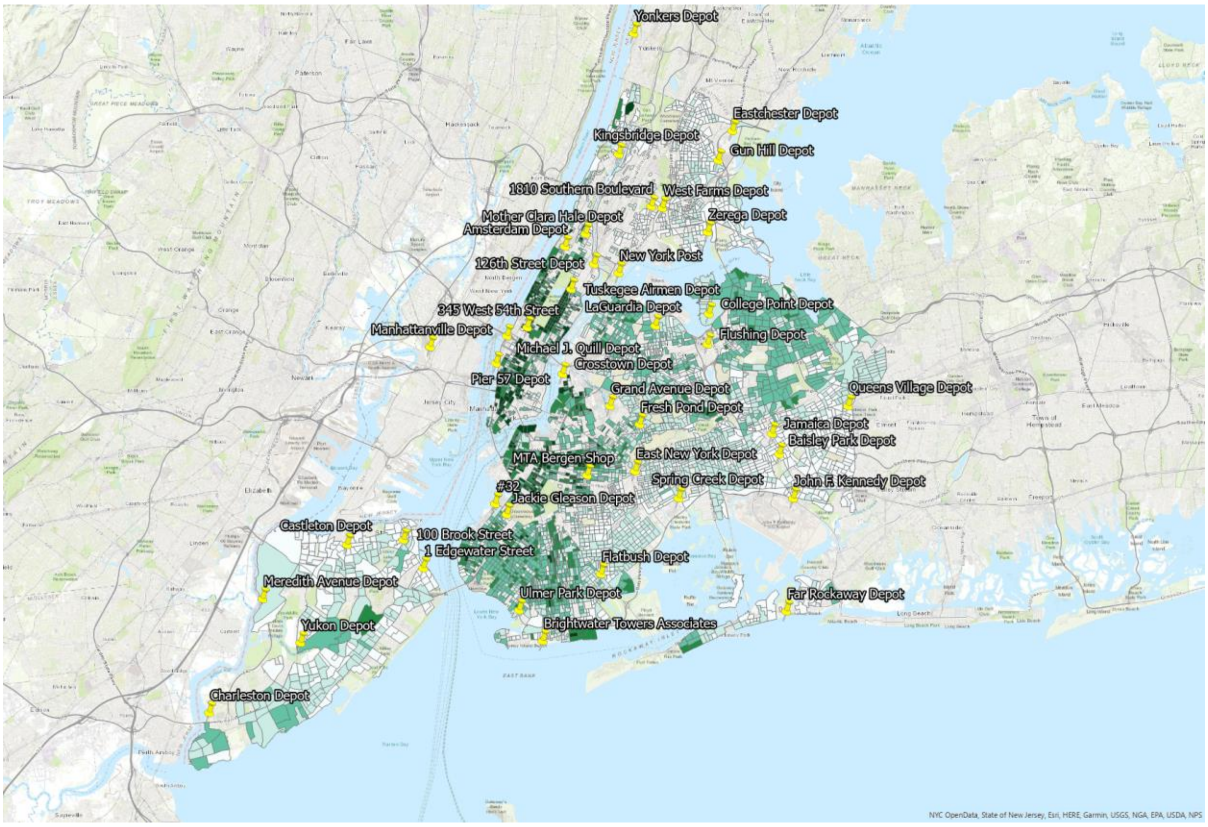

Figure 6: Mean Owner-Occupied Home Value in NYC Block Groups

5.2 Analyzing Pollution vs Distance

After performing geographically weighted regression on each of the five different types of pollution, I noticed that the models appeared to be close at first. However, after looking at the global R^2, I began to notice that the model was not perfect. In the scatter plot for black carbon emissions in Figure 7, there was a fair amount of scatter between the variables instead of a concentration. The global R^2 was only 0.14, which indicates a fairly weak correlation.

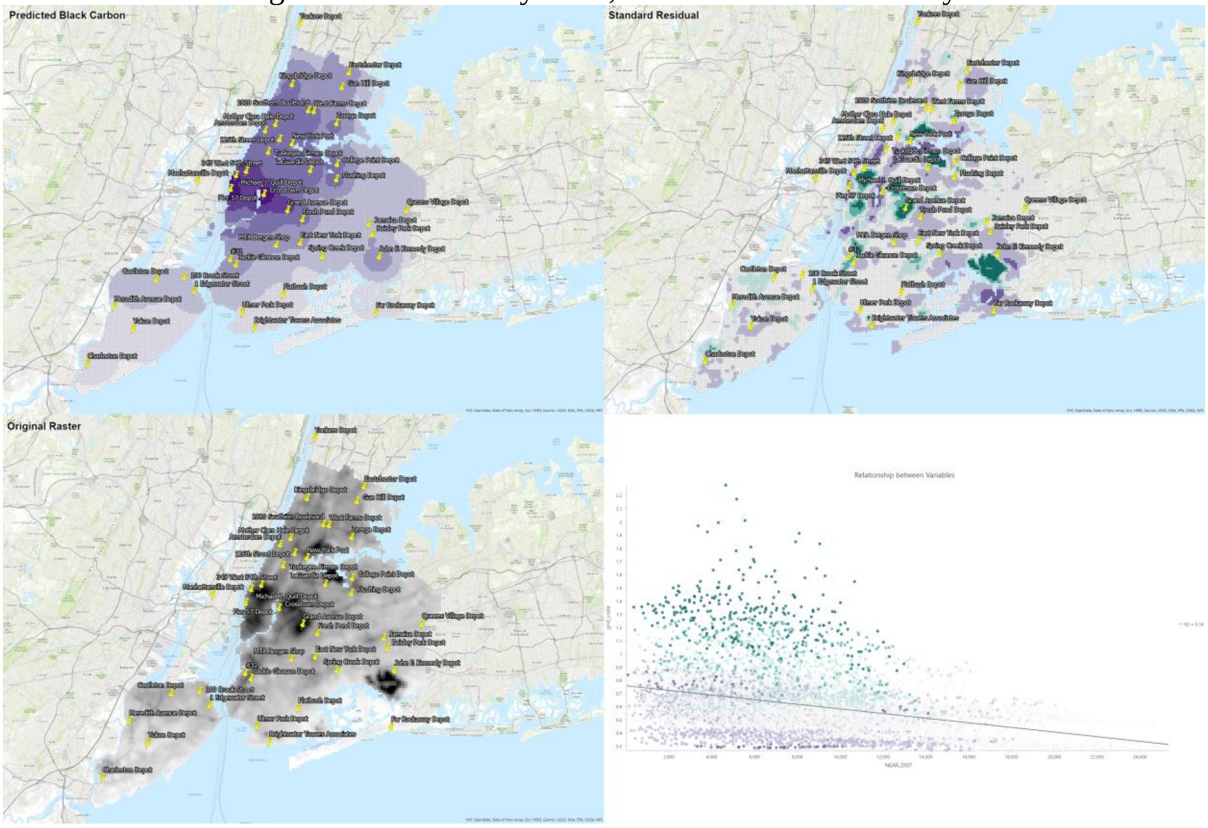

Figure 7: Black Carbon GWR Results

We can observe where the model's predictions were off by looking at the map of standardized residuals. In this map, the purple areas represent areas with a negative residual, which means that the predicted value was greater than the original value; thus, the model overestimated. On the other hand, green areas mark areas with positive residuals, where the predicted value was less than the original value, which means the model underestimated. By looking at Figure 7, we can see that the areas where the model under predicted matched with the hot spots in the initial raster. Likewise, the model appeared to overpredict in areas that a really low amount of black carbon. To me, this indicates that the model wasn't sensitive enough to extremes in the data.

Similar trends may be seen in the results for the other four pollution models, shown in Figures 8 through 11. All of the prediction models roughly portrayed the actual hot spots, but still tended to underpredict the values in said hot spots, while also overpredicting in areas with very low emissions. By looking at the scatter plots, we can also see that the correlations were all fairly weak. The global R^2 for all types of pollution were really low as seen in Table 1. This indicates that the relationships between different types of pollution and proximity to the nearest bus depot were all fairly weak by this model. This would suggest that either my results are inconclusive or that bus depots really don't have as significant an impact on air pollution as other possible sources.

Table 1. Global R^2 For Each Pollution GWR

Type	Value
BC	0.14
NO	0.09
NO2	0.18
FPM	0.12
Ozone	0.10

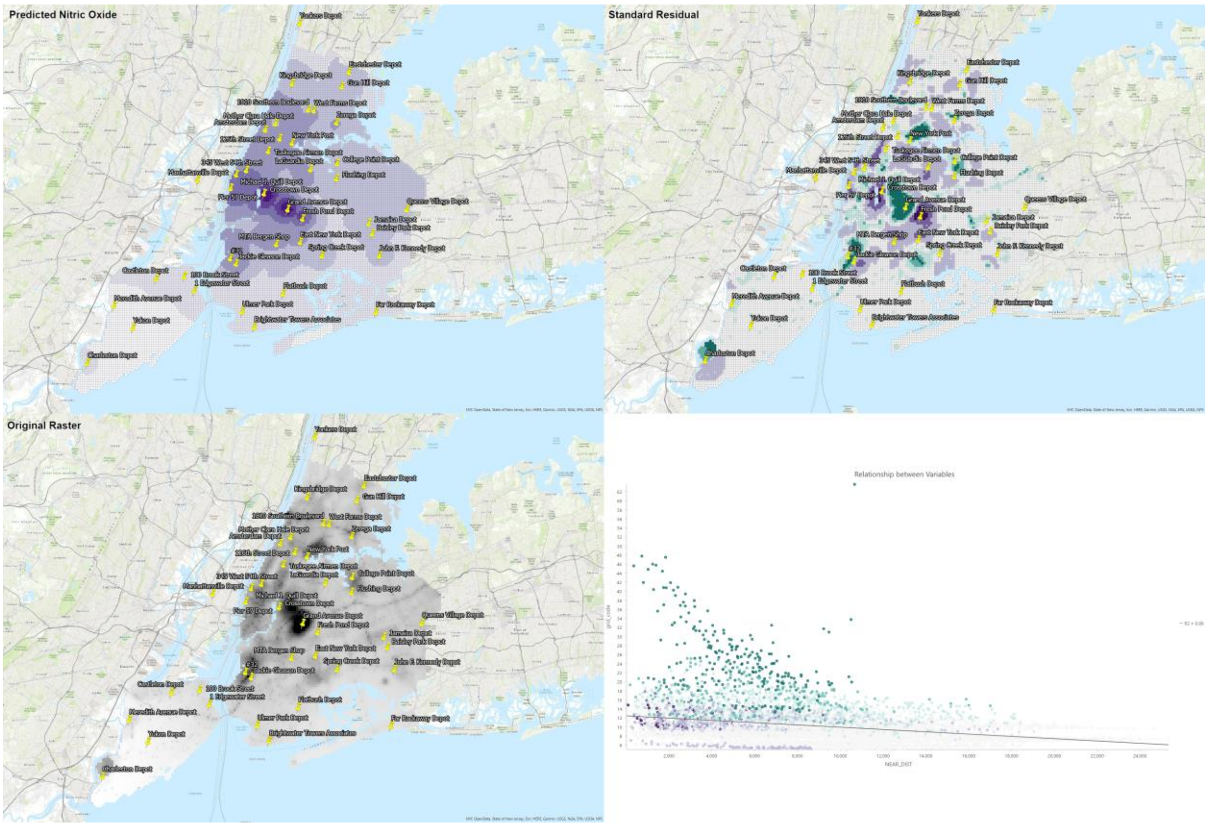

Figure 8: Nitric Oxide GWR Results

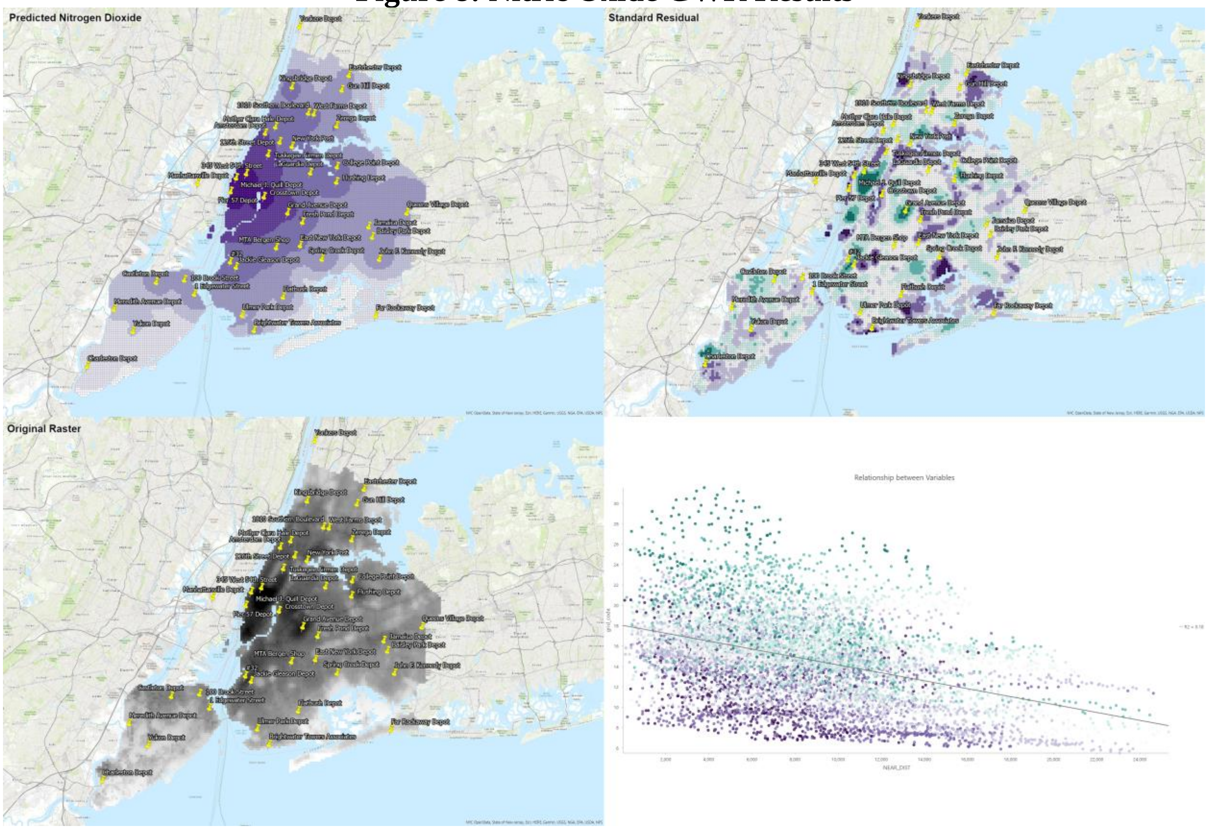

Figure 9: Nitrogen Dioxide GWR Results

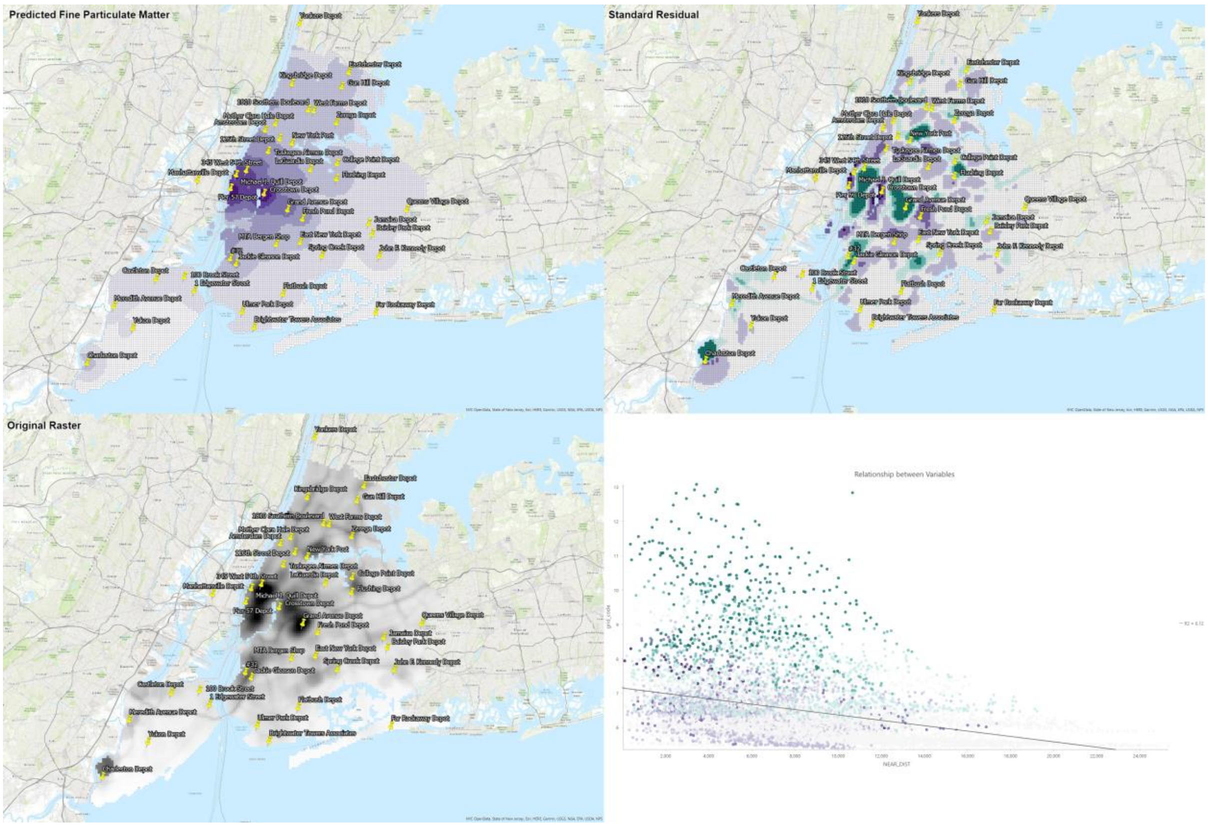

Figure 10: Fine Particulate Matter GWR Results

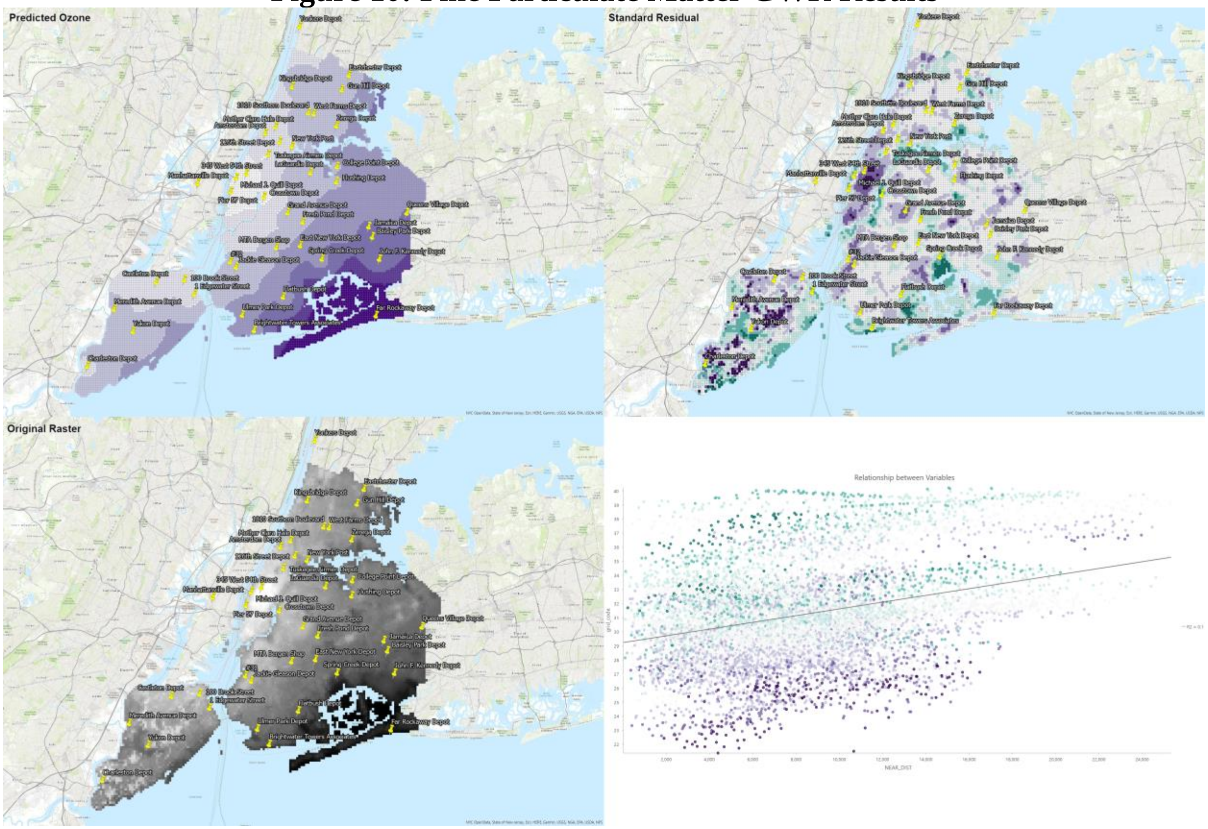

Figure 11: Ozone GWR Results

5.3 Modeling Land Value

For my final analysis, I performed geographically weighted regression on the mean home value for each census block group using all five types of pollution as well as distance from the nearest bus depot. In Figure 12, we can observe the calculated relations between each of the variables.

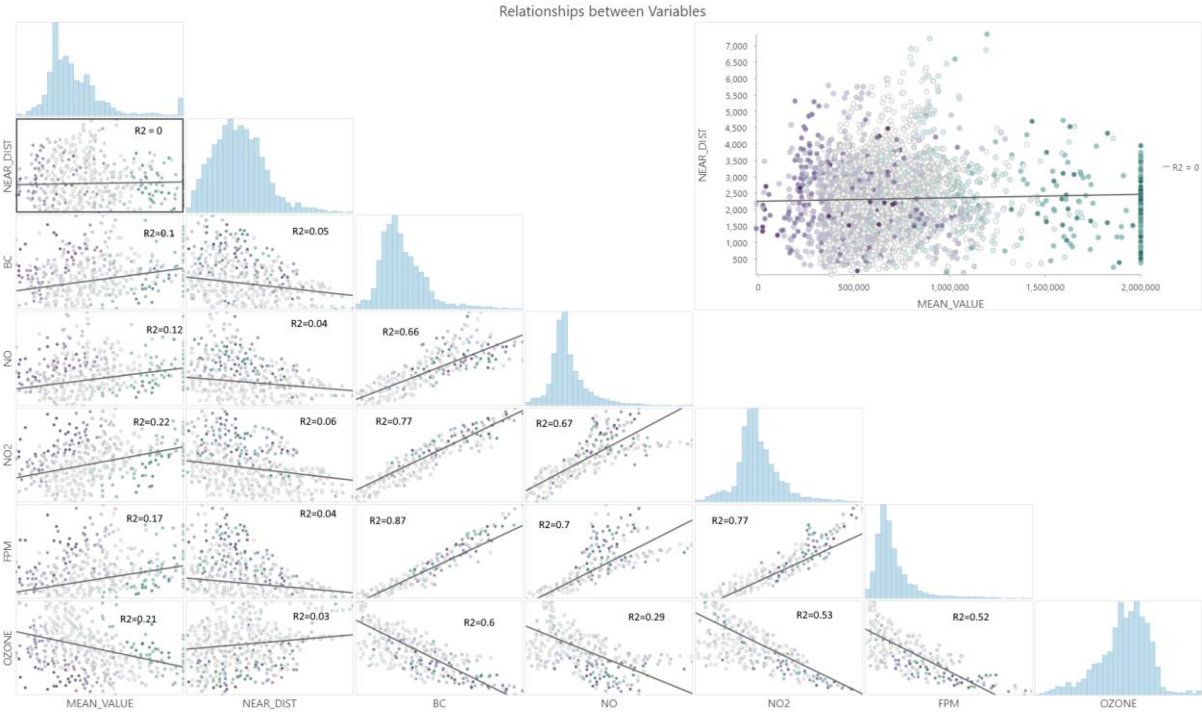

Figure 12: Mean Value GWR Variable Relations

Immediately, we can see that none of the different types of pollution seem to be closely related to mean home value. The relation directly between distance and value is the weakest with an R^2 of 0.0, indicating virtually no relation. The relation between distance and pollution isn't much stronger; however, the various types of pollution do seem to be related to each other, with ozone being the least related. I suspect this might have something to do with the initial being different when compared to the other four types of pollution.

In Figure 13, we can observe the standard residuals for each census block that was used for analysis. As some census block groups did not have either pollution data aggregated or because they had no data for mean home value, not every census block was used for this analysis and thus, do not appear on the map. However, we can still draw comparisons from the original data on mean value. A large portion of the census block groups are white or lightly colored, indicating that the predicted values were fairly close to the actual values. Similar to the pollution models, this model appeared to overestimate in areas that were less valuable, while underestimating in areas that were significantly more valuable, such as in Manhattan and northern Queens. There also do not appear to be any significant hot spots for extreme positive and negative residuals. Instead, the standard residuals seem to be more randomly distributed. This is good because it shows that there are not major areas where the model is consistently failing to accurately predict outcome.

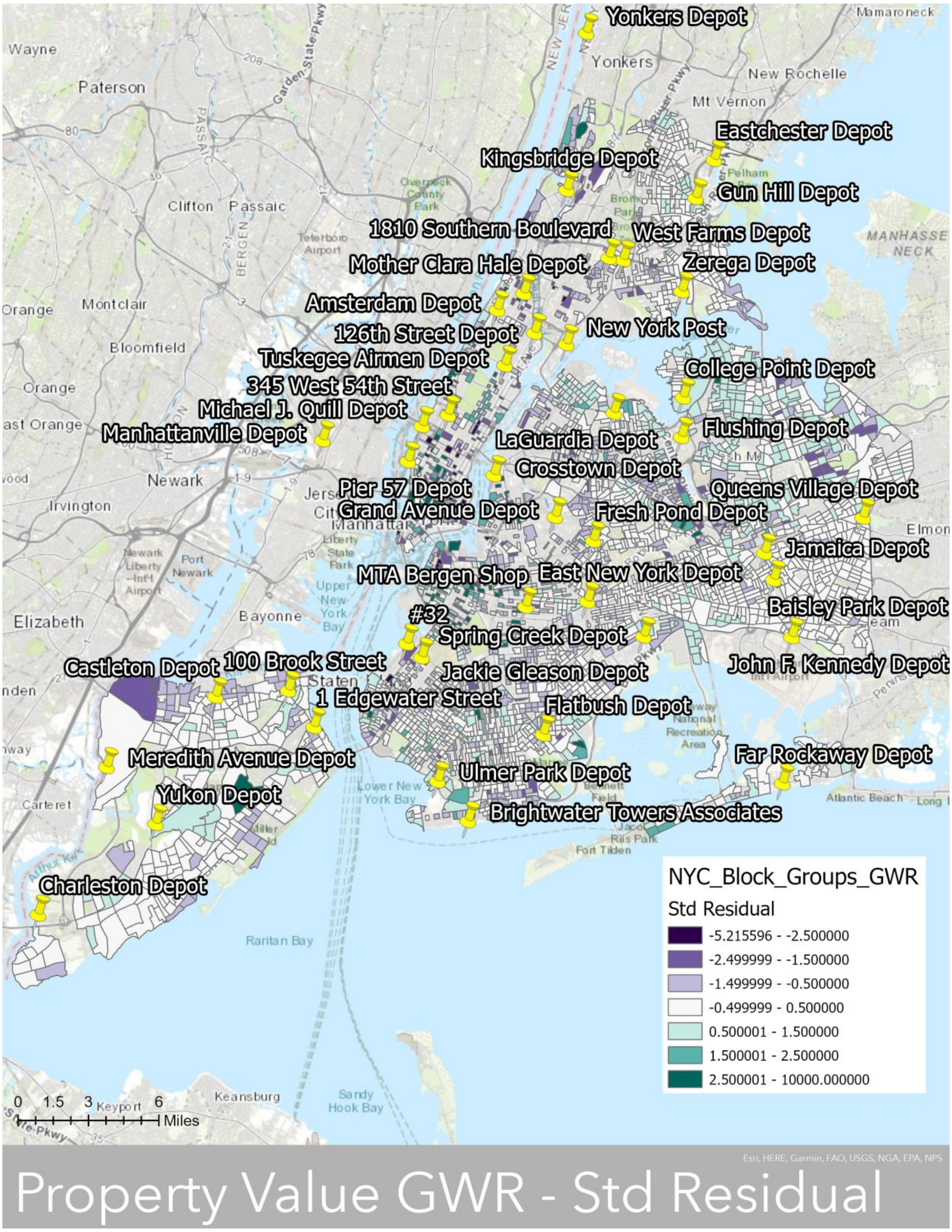

Figure 13: Standard Residual of Mean Value GWR

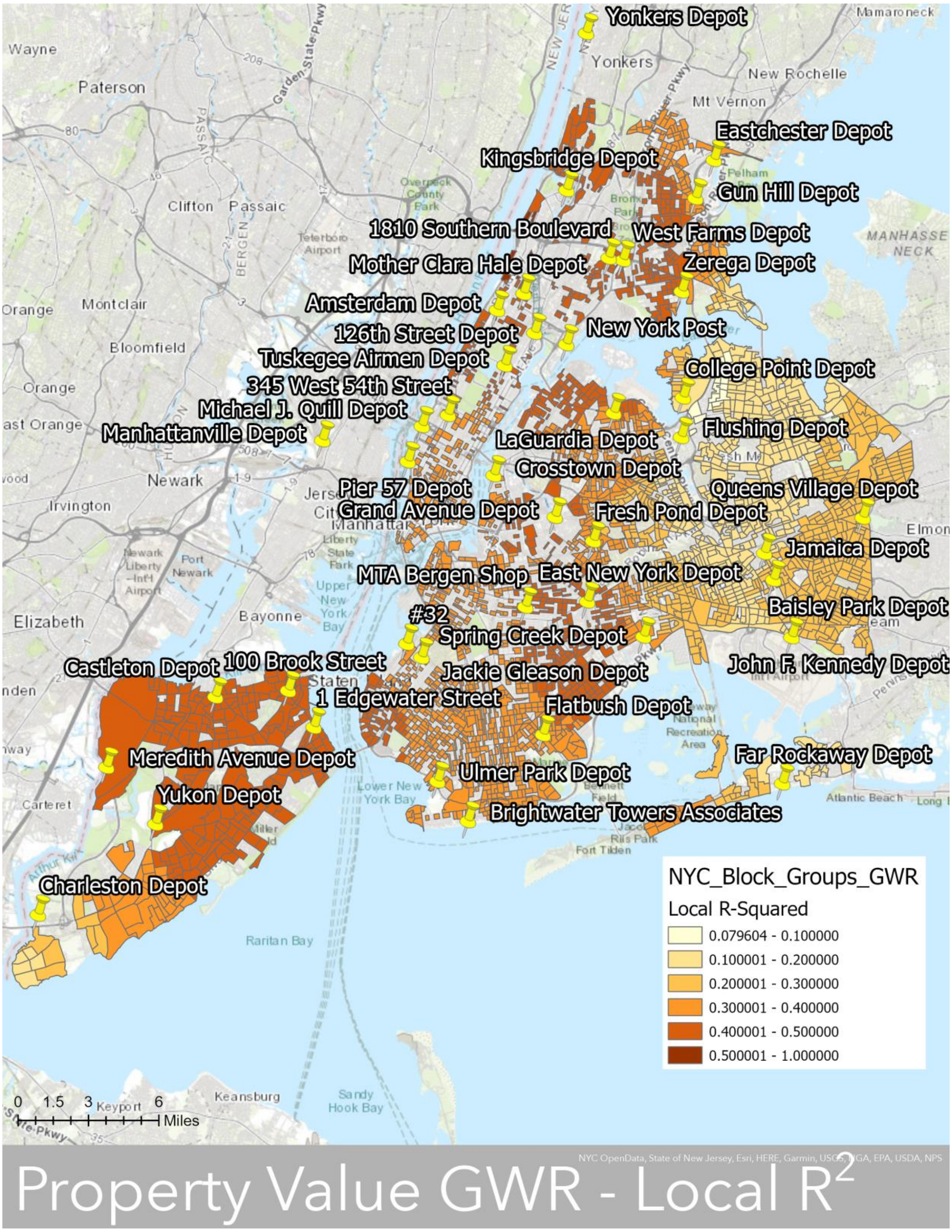

Figure 14: Local R^2 for Block Goups in Mean Value GWR

For this analysis, I have also mapped the local R^2 for each block group used. Rather than simply showing how the model fits for all of the data overall, it will instead highlight areas where the relations between the explanatory variables and mean value are stronger, thus showing where the model works better. In Figure 14, we can see that the relations were stronger in upper Manhattan and The Bronx as well in the northern part of Staten Island and a large portion of Brooklyn. Relations were weaker in Queens. However, even in areas where relations were stronger, the local R^2 rarely ever went above 0.5, indicating that relations were only moderate at best. This shows that the model is still not the best fit for the data.

Finally, I used this analysis to produce a raster model to predict mean value all across New York City. This raster model is meant to fill in the gaps left behind by the census tract data where there were no owner-occupied homes. This model can be seen in Figure 15. Here, we can see that the most valuable areas are in Manhattan while the least valuable areas are around Jamaica Bay, while Brooklyn and Queens appear to be moderately valuable. For the most part, this map roughly matches up with the recorded data we saw in Figure 6. Overall, the model looks decent; however, given that the relations between the variables were somewhat weak, it's likely that the model completely accurate.

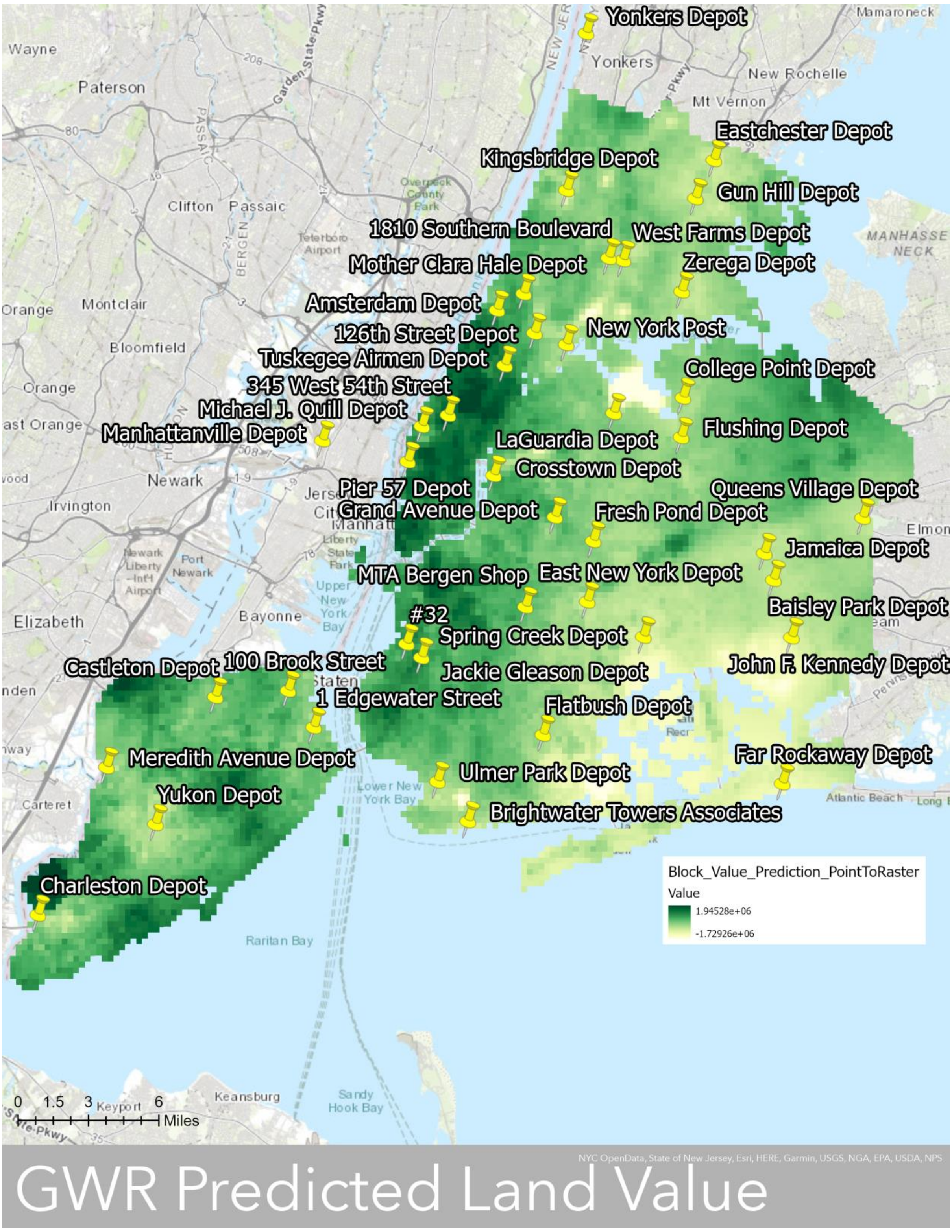

Figure 15: Mean Home Value Prediction Raster

6. Conclusion

This project's goal was to examine data on bus depots, air pollution, and land value, using geographically weighted regression. My main questions centered around whether bus depots had a major impact on air pollution in their surrounding environments and whether proximity to bus depots along with air pollution impacted land values. I initially predicated that there would be an increase in the levels of air pollution around bus depots and that this would negatively impact land values as a result. However, my analysis showed the relations between different types of air pollution and proximity to bus depots in New York City to be very weak. As such, the analysis is likely inconclusive or perhaps bus depots really don't have a significant impact as might be expected as compared to pollution from other sources. Likewise, pollution and bus depots appear to have little impact on land values in the New York City area based on this model. Once again, this may be due to an inconclusive analysis, or lack of real limited impact. It was interesting to note that the borough of Manhattan had both the highest amount of pollution and the highest property values. I suspect this is because it's the most densely populated area of the city and therefore would have more traffic.

As the results of this analysis were essentially inconclusive, I would wish to apply geographically weighted regression to relate other potential factors to pollution. For example, I now believe it would be more worthwhile to examine the effects of pollution by entire transit systems based on proximity to bus stops or subway stations in order to see if public transit would reduce pollution or increase it in certain areas. Traffic is another factor that I would like to study to see if transit routes reduce traffic in the areas that they serve, as well as traffic's effect on air pollution. It would also be worth analyzing how busy transit routes are as some routes are more heavily used than others.

Another major factor that I originally wished to study in this project but was unable to due to lack of data is noise pollution. When I first conceived of this project, I originally thought busses and by extension, bus depots, would have a much more significant impact on noise. This is still something that I would like to experiment with and study in the future. I imagine the lack of data is due to the fact that noise would be more difficult to collect data on and model since you would have to take into account the acoustics of the surrounding area. Thus, collecting data on sound is likely something that I would have to do myself using special audio equipment.

I would also like to experiment with different models as well as different neighbourhood selection methods and weighting schemes. Since this was the first time that I had modelled data using geographically weighted regression, I was essentially learning things as I went along. I had even learned things after I had performed my analysis that I believe could have made my work more accurate. For instance, solving issues with multi-collinearity when using multiple explanatory variables in my analysis. I am also wondering if using a number of neighbors as my bandwidth instead of a distance band would have produced more accurate results in addition to manually set intervals, as neighborhood selection is not something that I completely understand yet. Also, as with noise pollution, it may be worth it to go out and record pollution data myself as the original pollution data that I obtained was itself modeled based on the average annual recordings of pollution levels as opposed to raw data.

Another interest would be to analyze more cities outside of New York City, especially in the Pacific Northwest, such as Tacoma or Seattle. I originally tried to find data for Seattle or Tacoma, but was unsuccessful, which is all the more reason to record data on my own, or with a team. This kind of data collection would likely take several months, if not years. However, it might be done in collaboration with a local or state agency.

Overall, I believe that this project was a valuable leaning experience. I learned how to use regression analysis and, more specifically, geographically weighted regression. I also learned how to model data based on existing recorded factors and how to potentially apply it to others areas where I don't have data recorded. I have always had a passion for the environment as well as urban design and I hope to apply these new skills that I have learned for future projects that will make a positive difference.

References

1) Horvath, Ronald J. "Machine Space." *Geographical Review* 64, no. 2 (1974): 167-88. Accessed December 16, 2020. doi:10.2307/213809.
2) Pawlicki, Jerzy, Leszek Czapkowski, and Jacek Romańczyk. “Organizing a Bus Depot -- Selected Aspects.” *Transport Problems: An International Scientific Journal* 7, no. 1 (March 2012): 110–18.
3) Nylund, Nils-Olof, Kimmo Erkkilä, Maija Lappi, and Markku Ikonen. "Transit bus emission study: comparison of emissions from diesel and natural gas buses." In *IANGVA Conf, Buenos Aires*. (2004).
4) Patel, Molini M., James W. Quinn, Kyung Hwa Jung, Lori Hoepner, Diurka Diaz, Matthew Perzanowski, Andrew Rundle, Patrick L. Kinney, Frederica P. Perera, and Rachel L. Miller. "Traffic density and stationary sources of air pollution associated with wheeze, asthma, and immunoglobulin E from birth to age 5 years among New York City children." *Environmental research* 111, no. 8 (2011): 1222-1229.
5) Silverman, Robert A., and Kazuhiko Ito. "Age-related association of fine particles and ozone with severe acute asthma in New York City." *Journal of Allergy and Clinical Immunology* 125, no. 2 (2010): 367-373.
6) Kheirbek, Iyad, Jay Haney, Sharon Douglas, Kazuhiko Ito, and Thomas Matte. "The contribution of motor vehicle emissions to ambient fine particulate matter public health impacts in New York City: a health burden assessment." Environmental Health 15, no. 1 (2016): 1-14
7) Office of Budget and Policy, and Tina Hodges, Public Transportation's Role in Reducing Greenhouse Gas Emissions § (2010). https://www.transit.dot.gov/sites/fta.dot.gov/files/docs/PublicTransportationsRoleInRespondingToClimateChange2010.pdf.
8) Rivers, Nicholas, Soodeh Saberian, and Brandon Schaufele. "Public Transit and Air Pollution." *Available at SSRN 3049945* (2017).
9) Samaranayake, Samitha, Steven Glaser, David Holstius, Julien Monteil, Ken Tracton, Edmund Seto, and Alexandre Bayen. “Real-Time Estimation of Pollution Emissions and Dispersion from Highway Traffic.” *Computer-Aided Civil & Infrastructure Engineering* 29, no. 7 (August 2014): 546–58. doi:10.1111/mice.12078.
10) Apte, Joshua S., Kyle P. Messier, Shahzad Gani, Michael Brauer, Thomas W. Kirchstetter, Melissa M. Lunden, Julian D. Marshall, Christopher J. Portier, Roel CH Vermeulen, and Steven P. Hamburg. "High-resolution air pollution mapping with Google street view cars: exploiting big data." *Environmental science & technology* 51, no. 12 (2017): 6999-7008.
11) Herndon, Scott C., Joanne H. Shorter, Mark S. Zahniser, Joda Wormhoudt, David D. Nelson, Kenneth L. Demerjian, and Charles E. Kolb. "Real-time measurements of SO2, H2CO, and CH4 emissions from in-use curbside passenger buses in New York City using a chase vehicle." *Environmental science & technology* 39, no. 20 (2005): 7984-7990.
12) Shorter, Joanne H., Scott Herndon, Mark S. Zahniser, David D. Nelson, Joda Wormhoudt, Kenneth L. Demerjian, and Charles E. Kolb. "Real-time measurements of nitrogen oxide emissions from in-use New York City transit buses using a chase vehicle." *Environmental science & technology* 39, no. 20 (2005): 7991-8000.
13) Kheirbek, Iyad, Kazuhiko Ito, Richard Neitzel, Jung Kim, Sarah Johnson, Zev Ross, Holger Eisl, and Thomas Matte. "Spatial variation in environmental noise and air pollution in New York City." *Journal of Urban Health* 91, no. 3 (2014): 415-431.
14) Basner, Mathias, Wolfgang Babisch, Adrian Davis, Mark Brink, Charlotte Clark, Sabine Janssen, and Stephen Stansfeld. "Auditory and non-auditory effects of noise on health." *The Lancet* 383, no. 9925 (2014): 1325-1332.
15) Lall, Ramona, and George D. Thurston. "Identifying and quantifying transported vs. local sources of New York City PM2. 5 fine particulate matter air pollution." *Atmospheric Environment* 40 (2006): 333-346.
16) Civerolo, Kevin, Christian Hogrefe, Barry Lynn, Joyce Rosenthal, Jia-Yeong Ku, William Solecki, Jennifer Cox et al. "Estimating the effects of increased urbanization on surface meteorology and ozone concentrations in the New York City metropolitan region." *Atmospheric environment* 41, no. 9 (2007): 1803-1818.
17) Fricker, Ronald D., and Nicolas W. Hengartner. "Environmental equity and the distribution of toxic release inventory and other environmentally undesirable sites in metropolitan New York City." *Environmental and Ecological Statistics* 8, no. 1 (2001): 33-52.
18) Welsch, Heinz. "Environment and happiness: Valuation of air pollution using life satisfaction data." *Ecological economics* 58, no. 4 (2006): 801-813.
19) Mulley, Corinne, Liang Ma, Geoffrey Clifton, Barbara Yen, and Matthew Burke. "Residential property value impacts of proximity to transport infrastructure: An investigation of bus rapid transit and heavy rail networks in Brisbane, Australia." *Journal of Transport Geography* 54 (2016): 41-52.

20) Hui, Eddie CM, Chi Kwan Chau, Lilian Pun, and M. Y. Law. "Measuring the neighboring and environmental effects on residential property value: Using spatial weighting matrix." *Building and environment* 42, no. 6 (2007): 2333-2343.
21) Palmquist, Raymond B. "Property value models." *Handbook of environmental economics* 2 (2005): 763-819.
22) Haughwout, Andrew, James Orr, and David Bedoll. "The price of land in the New York metropolitan area." *Current Issues in Economics and Finance* 14, no. 3 (2008).
23) Mennis, Jeremy. "Mapping the results of geographically weighted regression." *The Cartographic Journal* 43, no. 2 (2006): 171-179.
24) Robinson, D. P., Christopher D. Lloyd, and Jennifer M. McKinley. "Increasing the accuracy of nitrogen dioxide (NO2) pollution mapping using geographically weighted regression (GWR) and geostatistics." *International Journal of Applied Earth Observation and Geoinformation* 21 (2013): 374-383.
25) Zhai, Liang, Shuang Li, Bin Zou, Huiyong Sang, Xin Fang, and Shan Xu. "An improved geographically weighted regression model for PM2. 5 concentration estimation in large areas." *Atmospheric Environment* 181 (2018): 145-154.
26) Zhang, Bo, Li, Weidong, Lownes, Nicholas, and Zhang, Chuanrong. "Estimating the Impacts of Proximity to Public Transportation on Residential Property Values: An Empirical Analysis for Hartford and Stamford Areas, Connecticut." *ISPRS International Journal of Geo-information.* 10, no. 2 (2021): 44-44
27) Liu, Yufeng, Steven Chien, Dawei Hu, Ning Wang, and Rui Zhang. “Developing an Extenics-Based Model for Evaluating Bus Transit System.” *Journal of Advanced Transportation*, November 19, 2020, 1–13. doi:10.1155/2020/8879664.
28) Jensen, Steen Solvang, Tim Larson, K. C. Deepti, and Joel D. Kaufman. "Modeling traffic air pollution in street canyons in New York City for intra-urban exposure assessment in the US Multi-Ethnic Study of atherosclerosis and air pollution." *Atmospheric Environment* 43, no. 30 (2009): 4544-4556.
29) Li, Qiong, Peng rui Bai, Yang Chen, and Xiao Wei. “Efficiency Evaluation of Bus Transport Operations Given Exogenous Environmental Factors.” *Journal of Advanced Transportation,* September 29, 2020, 1–13. doi:10.1155/2020/8899782.
30) Pemberton, Steve. "Optimising Melbourne's Bus Routes for Real-life Travel Patterns." *Case Studies on Transport Policy.* 8, no. 3 (2020): 1038-052.