Investigating the Relationship Between Construction Activity and Air Quality in New York City

Final Project for Civic Analytics GX-7003.001, Fall 2017

Evgeniya Bektasheva, Matt Dwyer, Lingyi Zhang
Center for Urban Science and Progress
New York University
NY, USA

Abstract: *Although construction activity is an important factor contributing to air pollution, there are very few studies examining the exact impact of construction on urban air. This research is aimed to fill in this gap by studying the correlation between PM2.5 concentration and the construction activities in surrounding neighborhoods.*

Keywords—construction, air pollution, PM2.5, New York

I. Overview

Air pollution in urban areas has become an increasingly important challenge in recent years. There are approximately 8.2 million people who live in New York City, and the number of tourists can double this number during certain times of year. City authorities are implementing various initiatives to improve air quality, with one such initiative being OneNYC, a "Plan for a Strong and Just City", that was released by Mayor de Blasio in 2015. The plan provides a set of initiatives and goals for improving resilience and equity in New York City, and among other goals, a plan to "achieve the cleanest air quality of any big U.S. city" by the year 2030 [1].

Understanding the extent of a pollution source's relationship with real measures of pollution could be helpful in predicting each source's impact. The analysis could lead to regulation that can specifically target pollution generators for the greatest air pollution mitigation.

II. Objectives and Social Impact of the Research

There are a number of factors that contribute to air quality in New York City. The research is examining some of these sources in more details and concentrates on pollutions coming from construction activity.

While construction sites pollution is difficult to measure, and such data is often not publicly available, most construction sites produce high levels of dust, which can be carried over large distances for a long period of time. Diesel engine exhaust from vehicles and heavy equipment, noxious vapors from oils, glues, thinners, paints, treated woods, plastics, cleaners and other hazardous chemicals that are widely used on construction sites also contribute to air pollution [2]. Therefore the research considers construction activity as a separate variable contributing to air quality in the city.

The major types of pollutants from construction are PM10 (particulate matter that is less than 10 microns in diameter, and is invisible to the naked eye) and PM 2.5 (fine particulate matter that is less than 2.5 microns in diameter). This research uses PM 2.5 data obtained from the monitoring sites in New York as a ground truth of measurement of air quality.

The results of the research findings could be used by city agencies primarily responsible for handling construction air complaints and construction regulation, such as the New York City Department of Environmental Protection (DEP) and the New York City Department of Buildings (DOB). Improvements in collaboration between DEP and DOB, improving communication and coordination between them could help increase efficiency in addressing construction air pollution problems.

Apart from that, exploring the sources and identifying the key factors that influence air pollution can provide a scientific basis for the formulation and effective implementation of social impact policies, and in particular, air pollution policies in urban areas [3].

III. Literature Review

A. Sources of Air Pollution in New York

A study on PM2.5 that was carried out in 2001 in Manhattan identified six source categories: regional transported sulfate, trans-continental desert dust, traffic, residual oil, ''local'' dust and World Trade Center fires pollution [4]. The New York City community air survey published in April 2017 found that areas with higher density of traffic, building, and heat and hot water boilers were also reported to have higher levels of pollutants, such as PM2.5, NO2, NO, and black carbon [5].

Traffic-related pollution is one of the major factors of pollution in NYC, and has been the subject of a number of bodies of research. It was revealed that traffic pollution is especially dangerous to populations living in neighborhoods with high densities of traffic [21]. Another major source of pollution is boilers in buildings. The density of oil-burning boilers combined with traffic density and temporality could explain as much as 84% of PM 2.5 variation in the winter 2008-2009 of New York City [6] .

Construction activity is also a source of air pollution [7]. Although New York City puts effort into controlling air quality issues from city construction, there is no air quality regulation that covers all construction stages similar to the regulation of noise from construction. Certain regulations cover only specific construction issues. Thus, Local Law 77 from 2003 and amendments specifically regulate the nonroad vehicles utilized in construction, such as limiting the use of diesel equipment to cleaner tiers and emissions control technology. There are regulations covering the demolition process, such as the Air Pollution Control Code, which outlines the prevention of airborne particulate matter, and prohibits the spraying of asbestos and insulation material [8].

It is difficult to find studies examining the correlation between air quality and construction activity in New York. An air quality analysis is performed on the construction phase of a project if the lead agency determines a significant amount of emissions are generated with respect to the total project emissions. Regarding construction activities that are short-term (less than two years, which is normally the case for most construction projects), such activities may either not be located near sensitive receptors, or as related to the construction of single buildings, no assessments were processed automatically [7]. The reason for the lack of analyses is that unlike fixed (or temporarily permanent) industrial and mobile transportation sources, the air pollution effects of construction activities are considered a temporary problem, and there is no regulatory requirements to monitor the ambient air pollutant levels associated with construction activities) [9].

Most countries, including US, do not report fugitive emissions from construction and specific emission factors for construction activities are uncertain [10]. Although it was confirmed that there is ample evidence that construction activities (such as land clearing, ground excavation, cut and fill operations and the construction of a particular facility itself) are an important source of particulate matter (PM) into the atmosphere and can have a substantial temporary impact on air quality, the emissions often vary substantially from day to day depending on the level of activity, the specific operations and the prevailing meteorological conditions making it difficult to assess the total contribution of such emissions to the air pollution levels of a city or region [10].

B. Pollution Effects on Public Health

A large number of studies have researched the dangerous effects of air pollution on public health. Air pollution relevant to construction activities contains both large particles and fine particles. Large particles mainly refer to the dust emissions during new facility building or the demolition activities. Fine particles come from construction equipment and transportation generated from construction needs, which mostly focus on PM2.5 and NO2 [7].

A study in 2016 confirmed that PM2.5 is a major harmful pollutant emitted from construction sites. It is persistent and stable in the air for some period of time, and can penetrate into the respiratory system [11]. In terms of the impact of air quality on public health, a report released by the NYC Department of Health details the health consequences of air pollution, estimating that between 2005 and 2007 PM2.5 has lead to over 3,100 deaths, along with thousands of hospitalizations for cardiovascular and respiratory diseases [12]. The report also estimates that efforts to reduce ambient PM2.5 concentrations have reduced a substantial number of health related incidents from the period of 2009 to 2011, such as 24% reductions in asthma related hospital visits, and 25% reductions in air quality related mortalities.

Dust and other air pollution from demolition and construction can impact the health and quality of life of people working and living nearby with some studies reporting an increment of mortality due to chronic obstructive pulmonary disease among construction workers[10]. Several studies also noted that certain populations are more vulnerable to pollution. Across 80 US metropolitan areas, exposure to ozone and fine particulates is higher in neighborhoods with greater proportions of “African Americans, Asian ethnic minorities, and poor households” [13]. Air pollution has also been shown to adversely impact elderly populations, as a study in Sao Paulo, Brazil, showed that the age group with the most significant increased risk of death were subjects over 65 years old [14].

C. Previously Utilized Methods

A study that analyzed the various sources of PM 2.5 in several cities and regions, including Connecticut and Massachusetts (US), Seoul (Korea), Genoa (Italy), Lecce (Italy), Ostrava (Czech Republic), Detroit and Chicago (US), used multivariate receptor models to identify the source apportionment of PM 2.5 [15]. The model is based on a natural science perspective and considers mainly chemical sources of air quality pollution, not taking into account socio-economic factors and various pollutants. A study that worked on degradation in urban air quality from construction activity and increased traffic used a regression model to calculate emissions from the road (including tail-pipe, brake and tire wear resuspension) and pollution arising from construction [10]. The model however lacked the addition of other regressors such as weather conditions.

Another statistical method used in an air quality study was a multilevel regression to measure the variation of regional compactness with exposure to poor air quality. When normalizing by population, the regression was able to show compact neighborhoods (measured through a sprawl index) as significant predictors of ozone levels [13].

Last, methods used in previous air quality studies relevant to the project include a poisson regression. The poisson regression was used to look at associations between public health issues (such as mortality) and air pollution, with adjustments made for seasonal variation and weather events. The regression model was then used to determine relative risk given estimated coefficients of air pollution variables for various mortality rates and age groups. Such a technique may be able to be applied to estimating risk of higher PM 2.5 levels based on construction [14].

IV. Data, Methods and Descriptive Analysis

A. Data Source Identifying and Cleaning

The project begun with identifying the data sources and extraction of the data. The data sources that are applicable to the research are listed in Appendix 1. It was assumed that cleaning the independent variables and dependent variables was necessary, along with aggregating each to the same geographical level so that they can be compared. To do this, the geographic level of zip code was chosen as its spatial resolution seemed small enough to isolate certain potential problem areas, yet be large enough to not contain too much error and be observable on a city-wide scale.

In the second step, the variables were studied for their relevance to the research. Since the research covers only the New York City area, the analysis and models that was built may not be used for other locations outside the City.

B. Construction Data

To measure construction, the duration of construction was chosen as a general indicator of the magnitude of construction activities, creating an assumption that longer projects are more likely to release greater PM2.5 concentrations. The study however does not account for further separation, such as indoor or outdoor construction activities. While indoor construction might not seem as directly related to air quality as outdoor construction, the related truck transportation of materials and debris, painting, and operating equipment can all contribute to air pollution [2]. With the lack of data to describe all these factors separately, project duration was the chosen variable to represent construction, which relies on two major assumptions. First, that the construction activities recorded in the Department of Buildings (DOB) permits dataset have equal pollution emissions on a daily basis, and second, that construction activity follows the records in the DOB data, starting on 'job starting day' and ending on 'permit expiration date.' These two assumptions may not reflect reality, as for instance 2635 cases were found to have a job start day later than the expiration day, meaning there were data entry errors, or construction activities did not follow the date in the records exactly. However, it is assumed that on aggregate the time between job start day and the construction permits expiration is a general indicator of the length of a project. Also, due to the absence of data on the scale of projects, it was assumed that construction activity had similar intensity in all cases.

C. Traffic, GHG and Boiler Data

Looking to traffic as a potential explainer of PM 2.5 rates, traffic data was pulled from the New York State Department of Transportation's traffic data viewer. This data came in the form of a points shapefile, with each point containing an estimated average daily traffic volume for a road. All the points were then spatially joined to a zip code boundary shapefile provided by the OpenData NYC website. The average of each traffic count was then calculated per zip code and then mapped in ArcMap.

Data for greenhouse gas (GHG) emissions from buildings was retrieved from the Local Law 84's energy data and disclosure, also made available via OpenData NYC. This data came in the form of Borough - Block - Lots (BBLs) with many geographic and energy related descriptors. The columns utilized in this analysis are greenhouse gas emissions, measured in metric tons of carbon dioxide equivalent (MtCO2e), and zip codes. The greenhouse gas emissions data contained outliers, which are assumed to be reporting errors as the standard deviation of 774,936 was much larger than the mean of 16,682 and the max value of 60,521,600 was much larger than the fourth quartile of only 1,026. To account for this disparity, only emissions data with a 95% quantile will be included in this analysis.

Oil boilers are one of the major PM 2.5 sources, so this variable was also considered for the research. This data was retrieved from the Mayor's Office of Long-Term Planning and Sustainability (OLTPS). In the analysis, annual total estimated consumption (High Estimate) in gallons was used as the indicator of oil boilers' contribution on air pollution. There were 8,048 buildings with oil boilers recorded, with 3,191 records having null values in the zipcode column. After extracting the zip code from the address location column, 8,045 entries were considered for further analysis.

D. PM2.5 Data

The PM2.5 data was considered in this study to be likely to exhibit strong positive spatial autocorrelation. This consideration was derived from two features of the PM2.5 data. First, due to fine articles diffusing in the air, the data was expected to be subjected to Tobler's Law, meaning near things are more related than distant things [16]. Second, autocorrelation could be present due to the method used to collect PM2.5 data. The data with the highest spatial resolution that could be found was provided by the New York City Community Air Survey, which records annually averaged concentrations in 42 United Hospital Fund (UHF) Neighborhoods. To geographically fit this data with the other pollution source data on the zip code level, the concentration value of each UHF was assigned into zip code areas using the Department of Health open data [20]. While this method was able to make the different datasets consistent, it contributed to the issue of higher spatial autocorrelation.

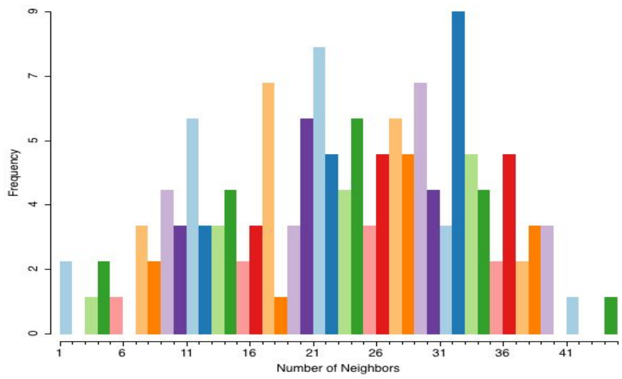

Fig. 1. Distribution of the number of neighbors for the weights matrix.

To test the autocorrelation for further spatial regression, choosing the most appropriate weights matrix is critical. The shapefile of zip code tabulation areas (ZCTAs) that was used did not interpolate areas between NYC's islands. Also, some ZCTAs have missing values, which were neglected in the shapefile used. As a result, several islands exist. In this situation, traditional weights matrix based on polygon contiguity (Queen and Rook) do not apply in the study. An alternative approach was to perform distance-based spatial weights instead [17]. Three commonly used distance-based spatial weights are inverse, distance bands, and K-nearest neighbors [18]. The latter two can be performed using GeoDa. The K-nearest neighbors method limits all points with the same number of nearest neighbors, which can not be used when doing spatial regression due to its asymmetric weights. The spatial regression is built on the algorithm of maximum likelihood, which requires weights to be symmetric [19].

Distance bands were found to be the most appropriate matrix in this research, not only because it generates symmetric weights, but also because of its diffusion manner which chooses all neighbors within a certain radius. In GeoDa, a threshold distance can be generated automatically, ensuring each unit has at least one neighbor to address the issue of islands. Since the distance bands measures the distance between two points, geopandas was used to produce the centroid of each polygon for building the matrix (Fig 1). Arc distance is employed with a threshold distance of 3.79913 miles, stored in the gwt format.

E. Univariate Analysis for all Variables and Descriptive Statistics

The spatio-temporal patterns were analysed in relation to construction activity in the City. In total, 390,595 construction activities were recorded in 2015. Most cases (over 70,000) covered the whole span of 2015. The cases that happened in the same zip code area were aggregated and calculated for the sum of days of all construction in one area. The sum of days of construction was considered as the indicator describing the level of impact of construction activities in a certain area.

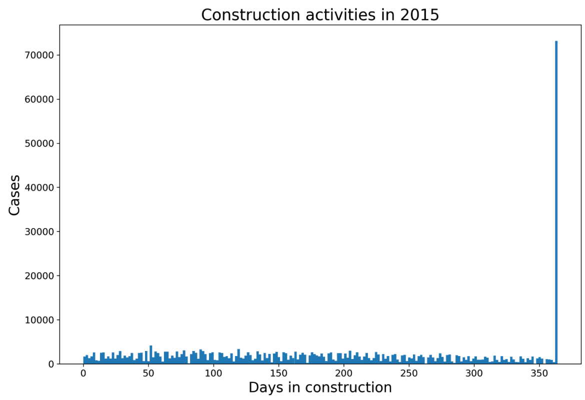

Fig. 2. Days of construction activity occurring in 2015.

The data was then mapped using the zip code shapefile (Fig. 3). The top five zip codes with the highest number of construction days are 10013, 10011, 10022, 10019, and 10001. These areas generally correspond to high construction activity areas in the city, such as Lower Manhattan, Hell's Kitchen, DUMBO in Brooklyn and Astoria in Queens. The five zip codes having the fewest days of construction are 10278, 10048, 10162, 11359, and 11005.

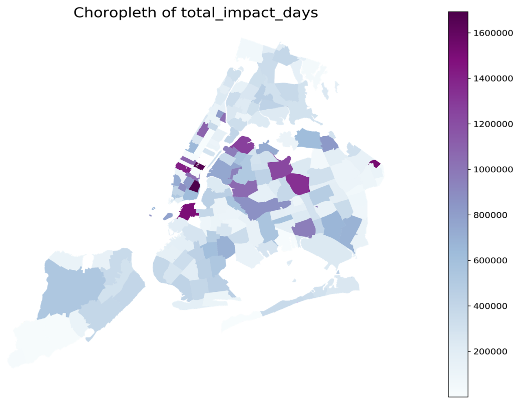

Fig. 3. Choropleth of the total impacting days in each zip code area.

Another variable relevant for the study was traffic volume. To plot the distribution of traffic volume counts, the dataset was cleaned to remove one extreme outlier with a traffic count of 103,324, which was interpreted as an error as it was almost 3 times larger than the next zip code with a count of 35,377. The dataset was further cleaned to exclude 79 zeroes, likely missing data due to lack of traffic counters. Descriptive statistics were performed on this variable as well, with New York City zip codes on average having traffic volume counts of 13,686 cars. The data appears to be relatively normally distributed (Fig 4).

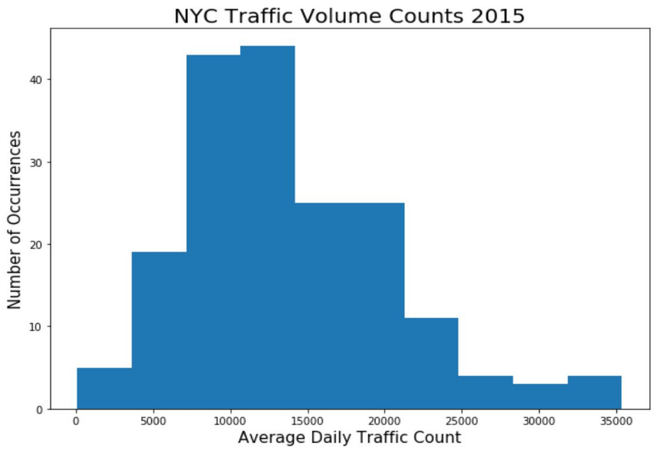

Fig.4. The distribution of traffic volume counts in New York City zip codes.

For the building greenhouse gas emissions data, about 500 upper outliers were removed, leaving 10,154 buildings. The distribution of the data was still extremely skewed right (Fig. 5), with a mean of 775 MtC02e, and an almost nearly as large standard deviation of 714.

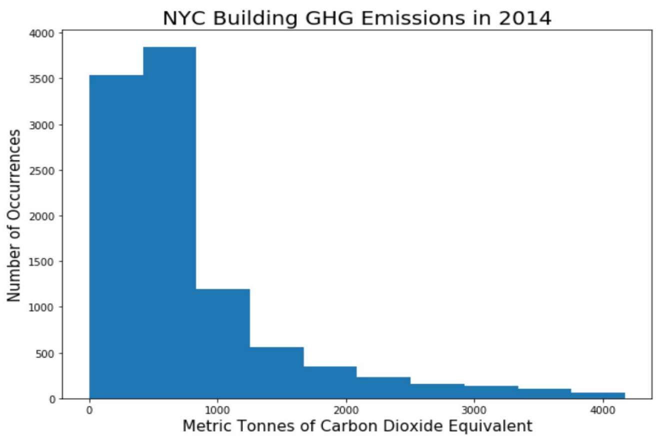

Fig. 5. The distribution of NYC building emissions in 2014.

The boilers data was also analysed for the research. For the 7883 buildings recorded with oil boilers, the annual sum of estimated consumption was calculated by zip code (Fig. 6). The distribution appears to be skewed right.

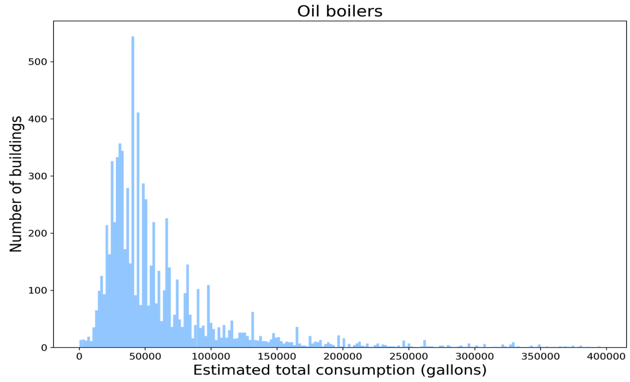

Fig. 6. Estimated total consumption by the oil boilers in 2015.

Finally, we considered two ways of analysing PM 2.5 data. Figure 7 below reflects the dataset which is the ground truth measure of urban air quality, represented by 34 monitoring sites.

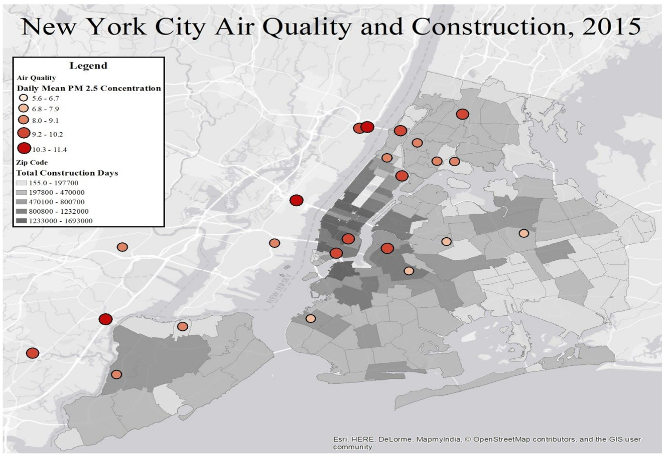

Fig. 7. Daily Mean Concentrations of PM 2.5

The zip code averages of daily mean PM2.5 concentrations ranged from 5.5 to 11.4 ug/m3 LC, averaging around 8.4. with each circle representing monitoring site, mapped against total construction days in 2015. The highest concentrations of both variables appear to be in Manhattan, while higher concentrations of PM 2.5 also appear across the Hudson River in New Jersey.

Obtaining valid PM 2.5 data was critical for the research, therefore the monitoring site data was backed up with data obtained from NYC Environment & Health Data Portal. This data was collected through New York City Community Air Survey, which was carried out on the United Hospital Fund (UHF) Neighborhoods (42) level. Based on the zip code definitions of New York City neighborhoods, the average PM 2.5 value was assigned to each zip code. Finally, 178 zip code areas were paired with PM2.5 values, with a standard deviation around 1.07 (Fig. 8). Generally higher concentrations are found in the north of the city.

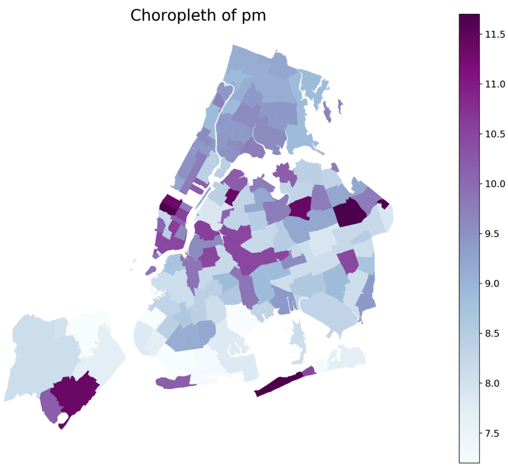

Fig. 8. Average PM2.5 concentrations of 2015 in New York City.

V. Results and Discussion

The final analysis sought to determine whether the number of days of construction activity positively correlate to high levels of PM 2.5 concentrations in 2015. This hypothesis was tested using a multivariate regression over the relevant variables that were considered for the research. The model used data from zip code level traffic, building emissions, and boiler emissions to find the extent to which each variable best predicts, and potentially contributes to PM 2.5 concentrations (Fig.9).

The autocorrelation test returned a Moran's I equal to 0.739603, which indicates a strong positive autocorrelation annually averaged density of PM2.5. As the local Moran's I shown in Fig. 10 indicates, high-high locations were found in Manhattan and northwestern Brooklyn.

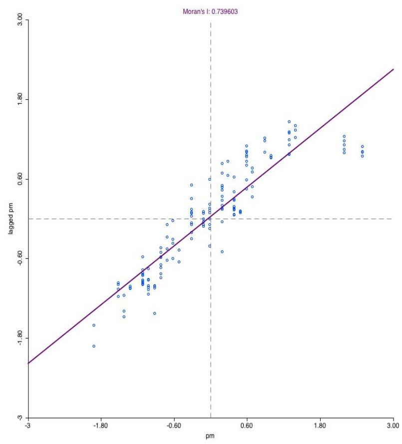

Fig. 9. Global Moran's I

These clusters correspond to the zip codes with high construction activity as previously found in Fig. 3. Meanwhile, low-low locations were located in the southeast of NYC, and again are generally corresponded to the zip codes with the low construction activity. Based on this result, positive spatial autocorrelation was found for PM2.5, so a spatially lagged dependent variable was introduced into the model.

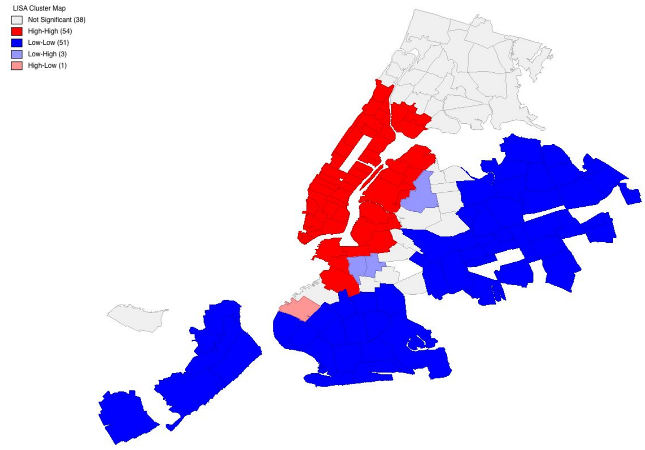

Fig. 10. Local Moran's I of PM2.5 concentrations.

As indicated in Table I, none of the independent variables are observed with high correlations. Boiler and construction have median correlation, meaning no serious multicollinearity.

TABLE I. Correlation matrix of variables

	Traffic	GHG	Boilers	Construction	PM2.5
Traffic	1.000000	-0.050510	0.394257	0.311007	0.223622
GHG	-0.050510	1.000000	0.094475	0.193270	0.379774
Boilers	0.394257	0.094475	1.000000	0.449864	0.425042
Construction	0.311007	0.193270	0.449864	1.000000	0.615641
PM2.5	0.223622	0.379774	0.425042	0.615641	1.000000

Initially the OLS regression on PM2.5 was performed against construction, GHG, oil boiler consumption, and traffic, without considering the effect of spatial dependence. With construction, GHG, and boilers having p-values smaller than 0.05, the R-square was around 0.4767, which suggested the model might not explain very much of the dependent variable. After introducing the spatial lagged dependent variable (W_pm), the overall performance of the models improved, with an R-squared of about 0.879. As traffic and boiler data gave insignificant results, feature selection was performed to reduce the complexity of the model. The feature selection step-backward was performed from model 1 to model 3. The

W_pm, construction, and GHG was kept in the final model (Table II).

TABLE II. Regression result of factors of PM2.5

Variables	OLS Model 1	Spatial Regression		
		Model 2	Model 3	Model 4
W_pm		0.927***	0.927***	0.931***
construction	0.478***	0.152***	0.154***	0.162***
GHG	0.272***	0.071**	0.070**	0.069**
boiler	0.177**	0.020	0.022	
traffic	0.019	0.008		
R^2	0.4767	0.8798	0.8797	0.8796

*p<0.1; **p<0.05; ***p<0.01

The results of the model show that PM2.5 concentrations have a high dependence in space. Compared to the OLS, the spatial lag models that control spatial dependence have improved the overall performance of the model. Construction activities and GHG emission in NYC have significant positive correlation with the PM2.5 concentrations. When the remaining variables are fixed, if the standardized construction increases one unit, the standardized PM2.5 will increase 0.162, meaning if the year-long construction activities cases increase by 23, the average pm concentration could rise roughly 0.094.

The final analysis of the model results is promising but quite counterintuitive. As noted in the literature review section above, there are very few studies related to construction activities in urban environment, and none of them specifically analysed the impact of construction on air quality. The results of the study, however, reveal that construction should be considered as an important factor contributing to PM2.5. The regression model revealed statistically significant result, i.e. overall model performance was quite high. The model showed statistical significance of each factor, and construction variable was the most significantly correlated. The correlation was stronger in the areas with high construction activity, and weaker in the areas with relatively few buildings that are being renovated or constructed, as the analysis in Fig. 3 revealed.

It could be possible that the impact of construction activities in urban areas had been underestimated before. The reason for that could be the fact that construction activity is usually associated with exhaust from vehicles and heavy equipment (that could fall under the traffic variable) or exhaust from oils (that could contribute to boiler emissions or GHG).

At the same time, it is also acknowledged that a multivariate regression that considered annual average data for only one year may not be the optimal model reflecting the influencing factors of the PM2.5 concentrations in New York. A main concern with the results of the model is that previous research showed that boiler and traffic emissions were supposed to be quite important pollution sources, but they are less significant in this study's model.

For further study, research into the effect of construction on air quality could be improved in several ways. For example, the analysis performed was on a zip code basis. Zip code level data could also be geographically linked to the more sparse air quality data by selecting only zip codes nearby to air quality monitoring sites, or, even more granularly, studying the buildings conducting large constructions within a certain radius of each sensor. Such an analysis would however require more sensors to be installed near large construction sites. Such opportunity was not available for our research, so we could not further separate the large construction projects.

Another method for further study is to find data pairing the construction durations with the PM2.5 concentrations in the exact same period. As described above, there is no data on the beginning and end of construction periods in New York, so certain assumptions had to be relied upon. If, however, data on exact days or months of construction periods could be obtained, such temporally specific data could improve the accuracy of the model.

Another aspect of the analysis to consider is the public perception of construction, looking at how people interact with pollution and construction. For such an analysis 311 data could be used to show construction-related complaints and could be further considered by zip codes or at a higher spatial resolution. While 311 data could be quite subjective and contain biases, it could bridge the divide between environmental measures and social impact.

References

[1] OneNYC: The Plan for a Strong and Just City. Retrieved 09-26-2017 from http://www.nyc.gov/html/onenyc/downloads/pdf/publications/OneNYC.pdf

[2] Gray, J. (2017, September 6). Pollution From Construction. Retrieved October 2, 2017, from

http://www.sustainablebuild.co.uk/pollutionfromconstruction.html

[3] Cheng, Z., Li, L., & Liu, J. (2017). Identifying the spatial effects and driving factors of urban PM 2.5 pollution in China. Ecological Indicators, 82, 61-75. doi:10.1016/j.ecolind.2017.06.043

[4] Lall, R., & Thurston, G. D. (2006). Identifying and quantifying transported vs. local sources of New York City

PM 2.5 fine particulate matter air pollution. Atmospheric Environment, 40, 333-346.

[5] The New York City Community Air Survey. Retrieved September 27, 2017, from:

http://www1.nyc.gov/assets/doh/downloads/pdf/environmental/comm-air-survey-08-15.pdf

[6] Clougherty, J. E., Kheirbek, I., Eisl, H. M., Ross, Z., Pezeshki, G., Gorczynski, J. E., . . . Matte, T. (2013). Intra-urban spatial variability in wintertime street-level concentrations of multiple combustion-related air pollutants: the New York City Community Air Survey (NYCCAS). Journal of Exposure Science and Environmental Epidemiology, 23(3), 232.

[7] CEQR Technical Manual, 2014 Edition (Revisions 4/27/2016). Retrieved October 22, 2017, from: http://www1.nyc.gov/site/oec/environmental-quality-review/technical-manual.page

[8] Title 24 of the Administrative Code of the City of New York. Chapter 1, Subchapter 6, Section 24-146)

[9] Air Quality Monitoring Study of Construction Activities between 69th and 87th Street on Second Avenue (Rep.). (2012, January 17). Retrieved October 23, 2017, from MTA Capital Construction website: http://web.mta.info/capital/sas_docs/AirStudyFinalReport-1-17-12.pdf

[10] Font, A., Baker, T., Mudway, I. S., Purdie, E., Dunster, C., & Fuller, G. W. (2014). Degradation in urban air quality from construction activity and increased traffic arising from a road widening scheme. Science of The Total Environment, 497-498, 123-132. doi:10.1016/j.scitotenv.2014.07.060

[11] Xing, Y.-F., Xu, Y.-H., Shi, M.-H., & Lian, Y.-X. (2016). The impact of PM2. 5 on the human respiratory system. Journal of thoracic disease, 8(1), E69.

[12] New York City Trends in Air Pollution and its Health Consequences (Rep.). (2013, September 26). Retrieved October 22, 2017, from New York City Department of Health website: https://www1.nyc.gov/assets/doh/downloads/pdf/environmental/air-quality-report-2013.pdf

[13] Schweitzer, L., & Zhou, J. (2010). Neighborhood Air Quality, Respiratory Health, and Vulnerable Populations in Compact and Sprawled Regions. Journal of the American Planning Association, 76(3), 363-371. doi:10.1080/01944363.2010.486623

[14] Gouveia, N. (2000). Time series analysis of air pollution and mortality: effects by cause, age and socioeconomic status. Journal of Epidemiology & Community Health, 54(10), 750-755. doi:10.1136/jech.54.10.750

[15] Lee, H. J., Gent, J. F., Leaderer, B. P., & Koutrakis, P. (2011). Spatial and temporal variability of fine particle composition and source types in five cities of Connecticut and Massachusetts. Science of The Total Environment, 409(11), 2133-2142. doi:10.1016/j.scitotenv.2011.02.025

[16]Tobler, W.R.: 'A computer movie simulating urban growth in the Detroit region', Economic geography, 1970, 46, (sup1), pp. 234-240

[17] Anselin, L.: 'An Introduction to Spatial Data Analysis, Distance-Based Spatial Weights ', 2017. https://geodacenter.github.io/workbook/4b_dist_weights/lab4b.html

[18] Anselin, L.: 'Spatial Weights' , 2002. http://www.dpi.inpe.br/gilberto/tutorials/software/geoda/tutorials/w8_weights.pdf

[19] GeoDa: 'Technical Questions' , 2017. Retrieved November 29, 2017, from: https://geodacenter.github.io/questions.html

[20] Converting index from neighborhood to zip code. Retrieved November 29, 2017 from: https://www.health.ny.gov/statistics/cancer/registry/appendix/neighborhoods.htm

[21] Patel, M. M., Quinn, J. W., Jung, K. H., Hoepner, L., Diaz, D., Perzanowski, M., . . . Miller, R. L. (2011). 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(8), 1222-1229.

Appendix 1.

Data sources

[Traffic] New York Department of Transportation. https://www.dot.ny.gov/tdv

[PM 2.5 Monitoring Sites] US Environmental Protection Agency. https://www.epa.gov/outdoor-air-quality-data/download-daily-data

[LL84] New York City. http://www.nyc.gov/html/gbee/html/plan/ll84_scores.shtml

[Oil Boilers - Detailed Fuel Consumption and Building Data] New York City.

https://data.cityofnewyork.us/Housing-Development/Oil-Boilers-Detailed-Fuel-Consumption-and-Building/jfzu-yy6n

[Construction - DOB Permit Issuance] New York City Department of Buildings.

https://data.cityofnewyork.us/Housing-Development/DOB-Permit-Issuance/ipu4-2q9a

[Zip Code] New York City. https://data.cityofnewyork.us/Business/Zip-Code-Boundaries/i8iw-xf4u/data

[PM 2.5 Concentrations by Community District] New York City Department of Health.

http://a816-dohbesp.nyc.gov/IndicatorPublic/VisualizationData.aspx?id=2023,719b87,91,Summarize