Journal of Public Health and Epidemiology

AJ ACADEMIC JOURNALS
expand your knowledge

Full Length Research Paper

Employing an Eigenfunction Eigendecomposition algorithm to cartographically and statistically delineate traffic-related carbon monoxide pollution in Hillsborough County, Florida

Jing Liu[1]*, Namit Choudhari[2], Brooke Yost[1] and Benjamin G. Jacob[3]

[1]Department of Epidemiology and Biostatistics, College of Public Health, University of South Florida, Tampa, Fl, USA.
[2]School of Geosciences, University of South Florida, Tampa, Fl, USA.
[3]Department of Global Health, College of Public Health, University of South Florida, Tampa, Fl, USA.

The traffic-related carbon monoxide (CO) pollution in Hillsborough County, Florida, has yet to be analyzed with high-resolution satellite data or algorithmic geo-spatiotemporal autocorrelation methods. This study aims to detect the association between traffic volume and CO pollution in Hillsborough County in 2022. It is hypothesized that daytime outdoor CO pollution positively correlates with traffic volume. High-resolution daily Giovanni remote sensing data of 1-degree spatial resolution from NASA Goddard Earth Sciences Data and Information Services Center was used to detect CO concentration on the roadways with annual average daily traffic (AADT) volumes larger than or equal to 80,000 in Hillsborough County. The results of AADT and CO concentrations indicated a clustering tendency. The Moran's Indices (*I*) of AADT was 0.956, and CO was 0.973. The results revealed a *non-homoscedastic* distribution of both AADT and CO concentrations. The performance of continuous data of CO and AADT was assessed by Pearson's correlation (r) to identify the strength and direction of CO concentration and AADT values. The results revealed a significant correlation between AADT and CO concentrations in this study with r=-0.119. The results also revealed an inversely proportional relationship; as AADT increased, daytime CO concentrations decreased. The clustered AADT and CO concentrations may account for biases in the correlation.

Key words: Asthma, annual average daily traffic, carbon monoxide pollution, remote sensing, Hillsborough County, Florida.

INTRODUCTION

Carbon monoxide (CO) is a harmful outdoor air pollutant mainly from burning fossil fuels, including motor vehicles and machines (National Aeronautics and Space Administration [NASA], n.d.; United States Environmental

Protection Agency [USEPA], 2022). Although CO is not a direct greenhouse gas (GHG), it aggravates the greenhouse effect by interacting with methane and carbon dioxide (NASA, n.d.; Voiland, 2015). Apart from this, CO pollution also poses considerable health risks to certain groups of people, including people with asthma, preexisting cardiopulmonary diseases, children and older adults (United States Consumer Product Safety Commission [USCPSC], n.d.; Raub, 1999; Kingsley et al., 2014; USEPA, 2014; USEPA, 2015; Yu et al., 2015; Prunicki et al., 2018; Ierodiakonou et al., 2019; Tang et al., 2019). Tang et al. (2019) found that increased personal exposure to outdoor CO pollution adversely impacts the cardiac autonomic control function of older residents in metropolitan areas and has a greater impact on cardiovascular health compared to young people. This is also because CO compounds in the ambient air can cause adverse health effects in humans at almost any concentration level (USCPSC, n.d.; Raub, 1999; USEPA, 2022). CO concentration at a low level may cause acute myocardial infarction in patients with heart diseases, and a concentration over 70 ppm may result in noticeable symptoms even in healthy people (USCPSC, n.d.; Raub, 1999). CO compounds are harmful to humans due to their nature to reduce the ability to carry oxygen in hemoglobin and thus increase the incidence of acute myocardial infarction in patients with heart diseases (Raub, 1999). Moreover, chronic exposure to CO at a low level may be more dangerous as it is difficult to predict the health effects of low-level CO exposure in individuals with ischemic heart disease (Raub, 1999). This indicates that chronic CO exposure may increase the risk of sudden death from arrhythmia in patients with coronary artery disease (Raub, 1999). In addition, CO exposure is also associated with increased levels of respiratory inflammation and thus results in asthma onset (Kingsley et al., 2014; USEPA, 2015; Yu et al., 2015; Prunicki et al., 2018; Ierodiakonou et al., 2019; USEPA, 2023b). There is abundant evidence showing that CO, together with other vehicle-related pollutants, is associated with genetic mutations that result in asthma in humans, especially in children (Yu et al., 2015; Prunicki et al., 2018; Ierodiakonou et al., 2019; USEPA, 2023b). Children are more vulnerable to air pollution than adults due to their underdeveloped respiratory systems and higher exposure frequencies from increased activity and breathing rates (Kingsley et al., 2014; USEPA, 2015). The smaller airways in children's respiratory systems pose even higher risks of exposure to pollutants relative to their size compared to adults (Kingsley et al., 2014). In the long term, children with developing respiratory systems may suffer from decreased lung function throughout their lifetime (Kingsley et al., 2014; USEPA, 2015). Furthermore, chronic exposure to air pollution may negatively impact children's cardiovascular health and neurobehavioral function (Kingsley et al., 2014). Considering the aforementioned risks, vehicle-induced CO pollution should be monitored and controlled.

The prevalence of CO pollution exposure in the United States is high (Kingsley et al., 2014; USEPA, 2014, 2023a, b). In 2017, 44% of the total CO emissions in the U.S. came from mobile sources, which was the primary source of CO emissions produced by humans (USEPA, 2023a). Additionally, nearly two-thirds of the traffic-induced emissions were created by vehicles on highways (USEPA, 2023a). It has been estimated that about 23 million (14%) people in the United States are susceptible to diseases caused by air pollution, and a quarter of them are children (USEPA, 2023b). Moreover, in 2009, more than 45 million Americans lived within 92 m of at least a highway, a railroad, or an airport (USEPA, 2014). Population trends indicate this number is increasing (USEPA, 2014). As children spend a large amount of time in school each year, often during peak traffic hours, they may be exposed to higher levels of traffic-related air pollution at school than at home (Kingsley et al., 2014; USEPA, 2014). According to the figure from a national assessment, from 2005 to 2006, approximately 6.4 million US children, or over 12% attended schools located within 250 m of major roadways, and as a result, they were exposed to high levels of traffic pollution (Kingsley et al., 2014). The effects of this exposure were particularly pronounced among minority and underprivileged children, although the extent of the impact varied depending on the region (Kingsley et al., 2014). Therefore, people in the U.S. are exposed to CO pollution, with specific populations having a higher risk (Kingsley et al., 2014; USEPA, 2014, 2023a, b).

Influential factors of CO pollution include traffic and weather conditions which can significantly affect pollutant concentrations (Kamiński et al., 2007; USEPA, 2014; Pan et al., 2016; Razavi-termeh et al., 2019; Hauptman et al., 2020; Abedian et al., 2021; Le et al., 2021; Njoku et al., 2022; Wang et al., 2021; Biswal et al., 2023). Traffic conditions that usually influence CO concentrations include vehicle volume, speed, and distance from sources of pollution, and at the same time, these factors interact with each other (Kamiński et al., 2007; USEPA, 2014; Pan et al., 2016; Tarigan et al., 2018; Razavi-termeh et al., 2019). Distance to the street contributed the most to the model testing association between traffic pollution and asthma onset (Razavi-termeh et al., 2019). According to Tarigan et al. (2018), the highest CO compounds aggregated at about seven meters from the traffic while dissipating to the background level at about two kilometers in their study. Moreover, Hauptman et al. (2020) found that living near major roadways increases respiratory health risks for children and pregnant women. Above 100 m away from major drive roads, with every 100 m increased in distance, children had about 30% fewer odds of having asthma onsets (Hauptman et al., 2020).

Traffic volume also plays an important role in creating CO pollution (USEPA, 2014; Pan et al., 2016; Wen et al., 2017; Le et al., 2021; Wang et al., 2021). Increased traffic volume generally leads to higher emissions (Pan et al.,

2016; USEPA, 2014). This can be demonstrated by some studies regarding the COVID lockdown. CO emissions decreased significantly during the lockdown due to reduced traffic intensity (Le et al., 2021; Wang et al., 2021; Orth and Russell, 2023). Furthermore, the concentration of CO emissions climaxes during peak hours in the mornings and evenings due to human activities, which also indicates the positive correlation between traffic volume and CO emission (USEPA, 2014; USEPA, 2015; Pan et al., 2016; Tarigan et al., 2018; Njoku et al., 2022). On the contrary, as traffic speed decreased, CO emissions increased (Pan et al., 2016; Abedian et al., 2021). Although pollutants normally can be reduced to background levels within 183 m away from the sources without considering weather conditions (USEPA, 2014), as the number of vehicles per hour increases, CO emissions can spread to over 500 m from the sources when traffic intensity and wind speed increase (Kamiński et al., 2007). The concentration of pollutants is also influenced by the type of vehicles (Orth and Russell, 2023), the design of roads, surrounding land use, and certain events such as congestion and acceleration (USEPA, 2014). In addition to traffic situations and human activities, weather conditions can also alter CO concentration (National Weather Service [NWS], n.d.; USEPA, 2014; Pan et al., 2016; Njoku et al., 2022). The pollutants tend to concentrate downwind on the road (USEPA, 2014). As wind speed increases, the CO dissipation rate increases (Pan et al., 2016; Njoku et al., 2022). According to Njoku et al. (2022), wind speed has more influence on CO concentration and dispersal than other weather parameters, such as humidity and temperature. The variations in CO concentration levels usually occur with the variation in weather and traffic conditions (Kamiński et al., 2007; USEPA, 2014; Pan et al., 2016; Razavi-termeh et al., 2019; Hauptman et al., 2020; Abedian et al., 2021; Le et al., 2021; Njoku et al., 2022; Wang et al., 2021; Biswal et al., 2023).

Although many studies have researched the associations between traffic and CO pollution, some issues within these studies might bias the results. In a study on about vehicle-related pollution, CO concentrations in the field were measured manually at a distance of 300, 500 and 600 m south of the identified center of the line source, involving factors such as traffic volume, wind direction, and speed, solar radiation intensity, and the map of Medan, Iran (Tarigan et al., 2018). Their results showed that the concentration of CO emissions peaked at seven meters from the source while decreasing to the bottom at two kilometers. However, this might be subjected to bias as only CO concentrations at the spots to the south of the street were measured (Tarigan et al., 2018). Another potential issue is that traffic volume and CO were measured on only one day; hence, the result could not represent the average level of CO emission and traffic conditions for the year. In a study conducted by Njoku et al. (2022), the authors built a program in ArcGIS Pro software to test the capability to forecast traffic-induced CO concentration. A significant positive association was found between vehicle emissions and CO pollution on the road. Although an empirical Bayesian Kriging regression prediction (EBKRP) model was applied in this study, CO was measured manually by researchers, and the selection of measuring spots was not explainable. Moreover, the time periods the authors measured during the empirical data collection only covered the peak hours and were not continuous. This may have created bias (spatial heteroscedasticity, that is, uneven variance multicollinearity, non-Gaussian error, and other violations of regression assumptions) in the results, which may have mis- specified the whole day calculated variable. Pan et al. (2016) applied the Gaussian dispersion model and puff integration model to predict traffic-induced pollution, including CO emissions, and their results showed a negative correlation between the speed of the vehicle flows and CO concentrations, while there existed a positive correlation between traffic volume and CO pollution. One of the problems with this study was that the authors estimated the CO concentrations utilizing a mathematical model that involved traffic-related parameters, including traffic speed and volume. This may have introduced bias into the relationship between the independent and the dependent variables when conducting the regression model. In particular, Gaussian diffusion models have limited applicability for relatively flat and homogeneous surfaces, reasonably steady and moderate to strong winds, moderately stable and unstable conditions, neutrally buoyant or slightly buoyant emissions, and relatively short distances from a source. In solving this issue, a spatial regression model should be employed to address the temporal-spatial features, such as associations in traffic volume and CO pollution data (Ali et al., 2021).

In the context of spatial regression analysis, several methods can be used to control for the statistical effects of spatial dependencies among estimated CO concentrations and traffic-related parameters. The maximum likelihood or Bayesian approaches account for spatial dependencies in a parametric framework, whereas recent spatial filtering approaches focus on nonparametrically removing geo-spatiotemporal autocorrelation. Spatial autocorrelation is the correlation among values of a single variable strictly attributable to their relatively close locational positions on a two-dimensional surface, introducing a deviation from the independent after observations assumption of classical statistics (Cliff and Ord, 1973; Anselin, 1988; Griffith, 2003).

In this paper, we propose a semiparametric spatial filtering approach that allows epidemiologists and other research collaborators to deal explicitly with (a) spatially lagged autoregressive models and (b) simultaneous autoregressive spatial models for quantitating geo-spatiotemporally dependent aggregation/non-aggregation-oriented propensities of CO signatures in high traffic volumes regions in Hillsborough County, Florida. As in

one non-parametric spatial filtering approach, a specific subset of eigenvectors from a transformed spatial link matrix is employed to capture dependencies among the disturbances of a spatial, regression, CO, and signature-traffic volume model (Jacob et al., 2013). However, the optimal subset in the proposed filtering model is identified more intuitively by an objective function that minimizes spatial autocorrelation rather than maximizes a model fit. The proposed objective function in a spatial autocorrelation model leads to a robust and smaller subset of selected eigenvectors (Griffith, 2003). Further, we assumed the application of the proposed eigenvector spatial filtering approach on georeferenced, multiple, mapped, high-traffic volume regions in Hillsborough County could demonstrate its feasibility, flexibility, and simplicity for prioritizing CO concentrated geolocations.

Vehicular congestion is a major problem in Hillsborough County and is managed by real-time control of traffic that requires accurate modeling and forecasting of traffic volumes. Traffic volume is a time series that has complex characteristics such as autocorrelation, trend and seasonality (Benjamin, 1986). Several linear and non-linear algorithmic modeling methods have been proposed to forecast traffic volume to support congestion control strategies in the literature (Akhtar and Moridpour, 2021). However, these methods focus on some environmental characteristics and ignore the latent non-zero autocorrelation in the data. Zero autocorrelation delineates geographic chaos (Griffith, 2003).

The present study attempts to develop non-zero geo-spatiotemporal autocorrelation models to predict CO concentrations at different mid-block sections of urban roads. The proportional share of vehicles and average traffic speed are considered inputs to the county-level prognosticative model. The traffic volume, speed and CO concentrations collected at different mid-block sections were analyzed. A good correlation was observed between average traffic speed, volume and CO concentrations. The assumption was that this study would show that classified traffic volume and average traffic speed in a mid-block could help explain the variance in CO levels significantly in a signature autocorrelation, forecast, and vulnerability model. In this study, both the ArcGIS Pro technique and remote-sensing data were adopted to minimize the potential latent geo-spatiotemporal autocorrelation bias. The schools, healthcare centers, and senior centers within a one-kilometer buffer from the highways with the most traffic volume values were identified by Google Maps and depicted by ArcGIS Pro software. The correlation between motor vehicle emissions and CO concentration levels was analyzed with SAS 9.4 software.

METHODS

Study site

Due to its subtropical location, Florida is more prone to be subjected to extreme weather events such as hurricanes and thunderstorms than the rest of the states in the U.S. (Florida Department of Health [FDOH], 2014; NWS, n.d.b). Located in the middle-west of Florida States, U.S., the size of Hillsborough County is about 1,020.3 square miles (United States Census Bureau [USCB], 2021) (Figure 1). Approximately 1,459,762 people were living in this county in the year 2020. The county median age was estimated as 37.9 years in 2021, which was younger than the median age in Florida of 42.8 years (USCB, 2021). At the beginning of 2022, the total number of registered vehicles and vessels in Hillsborough County was 1,359,866 (Florida Department of Highway Safety and Motor Vehicles [FLHSMV], 2022). In the year 2022, the average temperature was about 23.7°C, and the average precipitation reached about 1,268.476 mm in Hillsborough (NOAA National Centers for Environmental Information [NOAANCEI], 2023). As one of the major cities of Hillsborough County, Tampa had the highest average wind speed in March of 7.9 mph from 1983 through 2020 (Florida Climate Center [FCC], 2020).

Statistics processes / Pearson correlation coefficient test

SAS 9.4 Software was employed in conducting all the statistical analyses. A one-sample t-test was conducted to detect the normality and distribution of the AADT values. Two independent sample t-tests were conducted to examine the similarity of the lengths and the number of points on each road section between the two categories of AADT groups. A linear correlation test was applied to determine the Pearson correlation coefficient (PCC) values between AADT and CO concentrations. In statistics, the PCC is the ratio between the covariance of two variables and the product of their standard deviations; thus, it is essentially a normalized measurement of the covariance, such that the result always has a value of between -1 and 1 (Freedman et al., 2007). In this study, AADT values were classified into two groups: The high-AADT group and the medium-AADT group. The high level of AADT was defined as having AADT above 160,000 vehicles per day (VPD), and the medium-AADT group was defined as having AADT values from 80,000 to below 160,000. Based on this classification, two linear correlation tests were conducted between the two groups of AADT variables and the corresponding CO concentrations.

Identifying essential buildings

First, ArcGIS Pro 2.9 was utilized to generate one-kilometer buffers surrounding the driveways with annual average daily traffic (AADT) values greater than or equal to 80,000 VPD with the geoprocessing tool. Second, Google Maps was employed to search for essential buildings located within the buffers on July 7th, 2023. After all the locations of the essential buildings were identified, the coordinates were also documented and imported into ArcGIS Pro 2.9. A map with all buildings with pairwise buffers was subsequently generated.

Variables / AADT values and CO concentrations

The independent variable is the AADT values, and the dependent variable is the CO concentrations. The AADT approximates the number of motor vehicles going through a section of a roadway on average over the year 2022 (Datagov, 2022). The values were calculated through standardized formulations provided by the Federal Highway Administration (Federal Highway Administration [FHA], 2018). Using AADT can minimize bias when calculating the daily traffic volume (FHA, 2018). The dataset of AADT values in the year 2022 was obtained from the Florida Department of Transportation on July 5, 2023, from the Florida Department of Transportation governmental website (Florida Department of
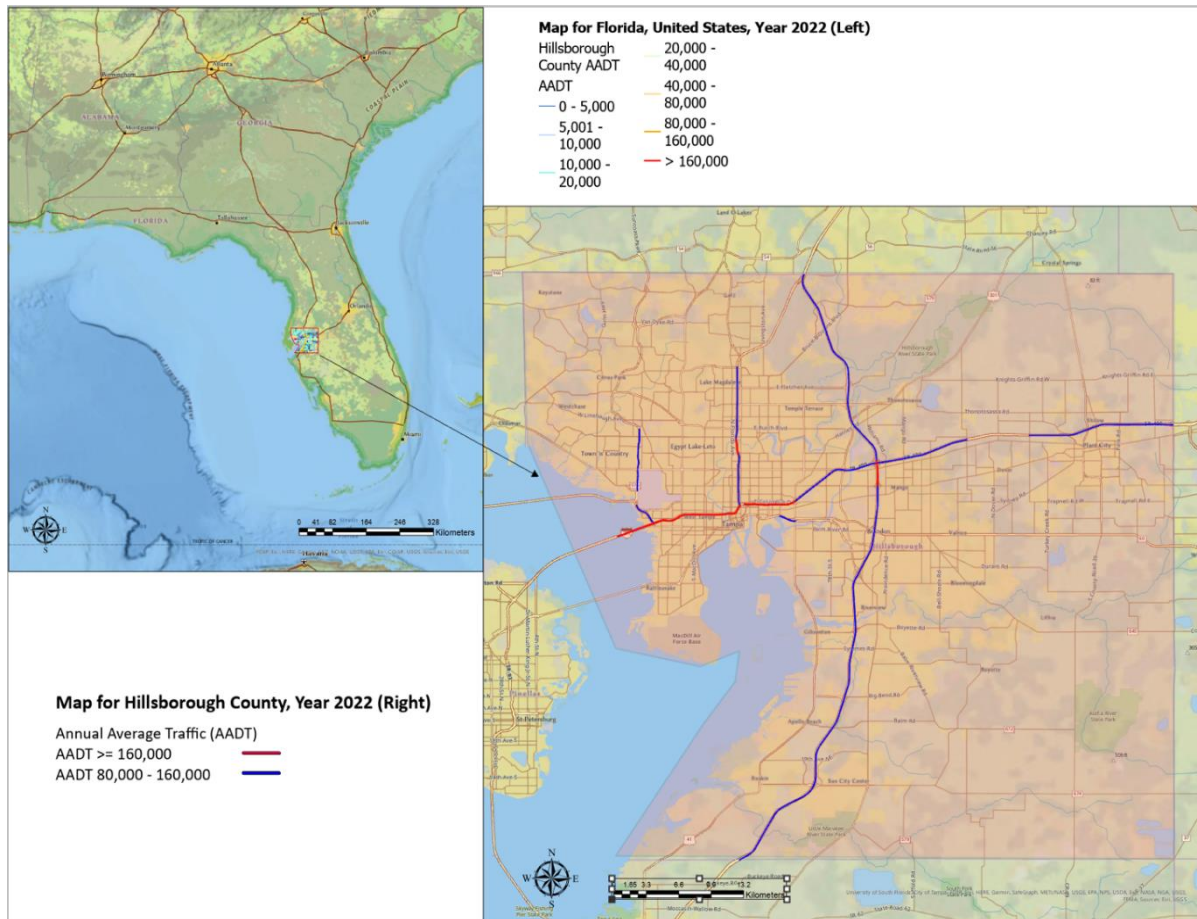
Figure 1. Map of Hillsborough County within Florida, USA, made by ArcGIS Pro.
Source: Authors.

Transportation [FDOT], 2022). The measuring points on each section of the roadway were added manually. First, the minimum (min=379.9 m) and maximum (min=10418.3 m) lengths of road sections were identified. SAS 9.4 software was used to calculate the number of points added to the sections of the roadways. The number of capture points was calculated by dividing the length of the road section by 300, and the result was rounded down to get the number of points in this section. The formula is shown as follows:

$$Number\ of\ points = \frac{Length}{300}$$

The dependent variable was the CO concentration. The daytime CO concentration values of 2022 were downloaded on July 9, 2023, from NASA Goddard Earth Sciences Data and Information Services Center (GES DISC) (Meyer, 2022). The values on the measuring points were obtained in terms of high-resolution daily Giovanni remote sensing data of 1-degree spatial resolution. The correlation between AADT and CO concentrations was examined by PCC (r) with SAS 9.4 software. The AADT values were stratified into two categories, which were the high-AADT (AADT ≥ 160,000 VPD) and medium-AADT (80,000 ≤ AADT < 160,000 VPD) group (Figure 2). The correlations between the CO concentrations and the two categories of AADT values were analyzed with PCC (r) with SAS

9.4 software.

Eigendecomposition

The assumption for spatial independence was tested for the traffic volume and CO sampled; the Pearson product-moment correlation coefficient (that is, Moran's *I*) was employed for observations. Moran coefficient is an index of spatial autocorrelation involving the computation of cross-products of mean-adjusted values that are geographic neighbors (that is, covariations) that ranges from roughly (-1, -0.5) 0 to nearly 0 for negative, and nearly 0 to approximately 1 for positive, spatial autocorrelation, with an expected value of -1/(n-1) for zero spatial autocorrelation, where n denotes the number of areal units (Griffith, 2003).

Moran's *I* was employed as a diagnostic tool for quantitating model misspecifications, spatial non-homoscedasticity (that is, uneven variance), and outliers in the remotely sensed traffic volume and CO sampled parameter estimators. Homoscedasticity describes a situation in which the error term (that is, the "noise" or random disturbance in the relationship between the independent variables and the dependent variable) is the same across all values of the independent variables (Freedman et al., 2007). The frequency dataset was stratified into georeferenced groups of traffic volume and CO sampled, with proportions based on their mid-block occurrence abundance and distribution. Likewise, Moran's *I* was
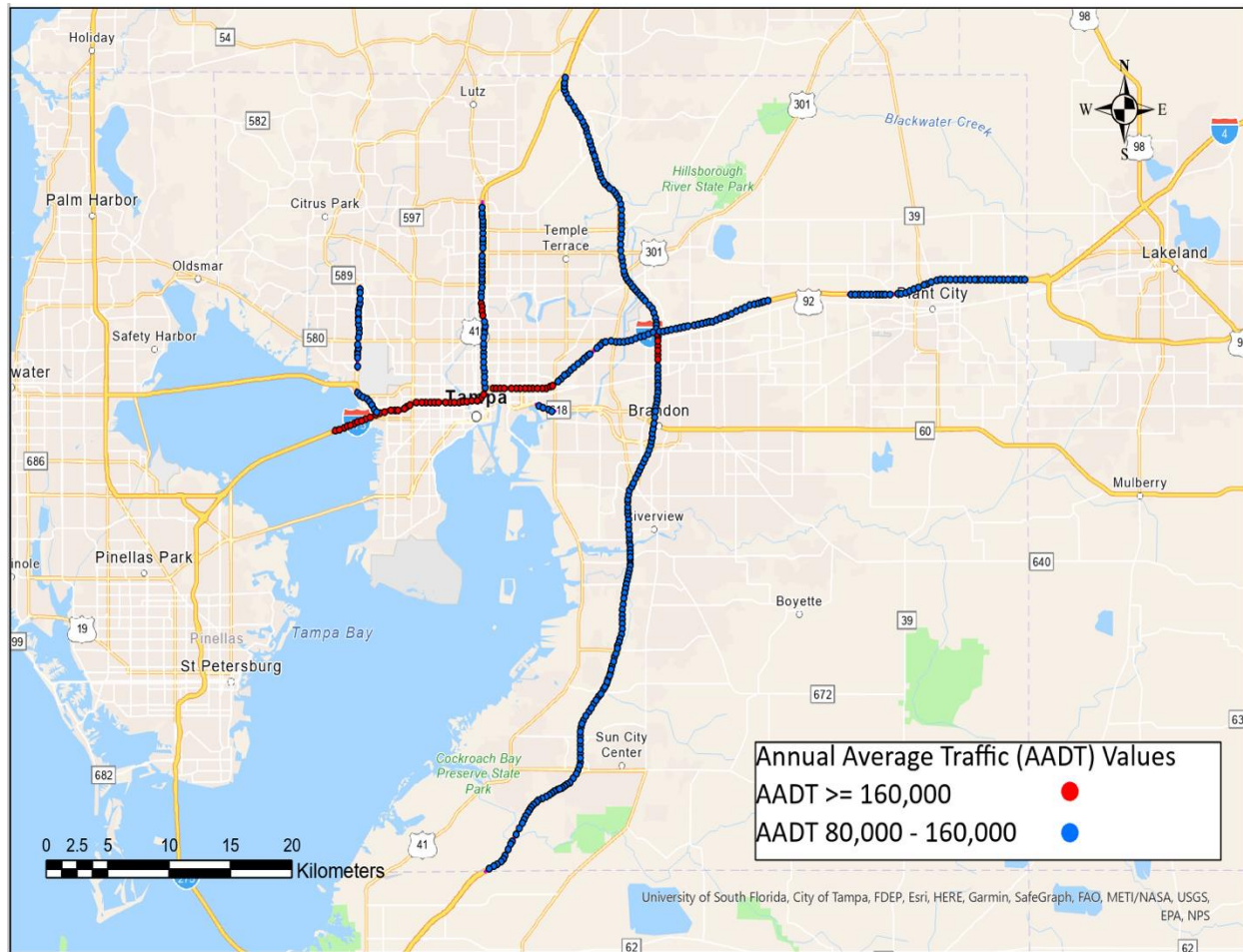
Figure 2. Map of measuring spots with AADT variances in Hillsborough County, FL, 2022.
Source: Authors.

employed to determine if the dependent variables were clustered or randomly distributed within a geographic space in Hillsborough County.

ArcGIS PRO was employed to generate Moran's *I* by computing the cross mean of Euclidean inter-site distances between stratified traffic volume and CO sampled, mid-block measured explanatory, discrete integer values that were geographic neighbors. In the first step in Moran's I analysis, "*neighboring" polygons* (that is, contiguous polygons, polygons within a certain Euclidean distance) was defined (Cressie, 1993). The LAGDISTANCE OPTION indicated the neighborhood size, which was important in the computation of the autocorrelation index for quantitating clustering propensities in the sampled variables. It is of note that lag distance in this research was dependent on the sampled county-level traffic volume and CO sampled, parameterized estimator sample dataset. The goal was to create a variogram that invariably provided optimal estimates of spatial dependence for the underlying stochastic process within the dataset.

The compute statement allowed the averaging of binary spatial weights within the autocorrelation statistical process needed for the construction of Moran's *I* coefficient (an equivalent of regression slope for Moran's scatter plot). Using the values from LAGDISTANCE and MAXLAGS, the traffic volume and CO sampled frequency model in ArcGIS Pro without the Novariogram option in order to compute the empirical semivariogram was constructed. A variogram is often defined as a measure of spatial variability (Griffith, 2003). The strategy was that by sampling stratified traffic volume and CO sampled capture points close to each other, this would produce typically similar outcomes compared to sampling for the capture points separated by larger distances in geographic space. Here, the variogram distance measured the degree of dissimilarity $\gamma(h)$ between the sampled stratified traffic volume and CO sampled data separated by a class of vectors h. If z (xi) and z(xj + h) were pairs of exploratory georeferenced, traffic volume and CO observational samples lying within a given class of distance and direction, where N(h) was the number of data pairs within an urban commercial land cover class. Subsequently, the experimental semivariogram was defined in ArcGIS Pro as the average squared difference between the components of the sampled, the traffic volume and CO sampled, stratified data pairs in geographic space employing the following equation:

$$\gamma h = 12Nh\sum i = 1Nhzx - xi + h2]\tag{1}$$

This spatial variability measure is a semivariogram (Cressie, 1993). The study interpolated between the sample variogram, traffic volume and CO sampled, explanatory, time series, and dependent estimators. The variance of the entire dataset was re-defined as the sill and the distance at which the model semivariogram met the data set variance, which, in this research, we defined as the range.

A robust version of the semivariance was requested with the ROBUST option in the COMPUTE statement. ArcGIS Pro rendered a plot showing both the classical and the robust empirical semivariograms. The plot option to specify different instances of plots was featured in the empirical semivariogram. In addition, the autocorrelation Moran's *I* statistics was generated under the assumption of randomization using binary weights. The output from the requested autocorrelation predictive, probabilistic, geo-spatiotemporal analysis included the observed, computed Geary's c coefficients. The finely tabulated expected value and standard deviation for each sampled stratified, traffic volume and CO sampled, explanatory, covariate the corresponding Z score, and the p-value were calculated in the Pr >j Z j column. The low p-values suggested non-zero autocorrelation for both statistics types. Note that a two-sided p-value was generated, which was based on the probability that the observed traffic volume and CO sampled frequency coefficients lay farther away from j Z j on either side of the coefficient's expected value, that is, lower than Z or higher than Z. The sign of Z for both Moran's *I* and Geary's c coefficients can indicate latent positive or negative geo-spatiotemporal autocorrelation (Griffith, 2003). The output randomization estimates from the stratified traffic volume and CO sampled, autocorrelation, and frequency model were then evaluated in a spatial error (SE) model. An autoregressive model was employed whereby a geosampled, temporally dependent, socioeconomic stratified variable, Y, as a function of nearby sampled traffic volume and CO sampled, frequency Y values [that is, an autoregressive response (AR) or spatial linear (SL) specification] and/or the residuals of Y as a function of nearby Y residuals (that is, an AR or SE specification). Distance between frequency-sampled, georeferenced stratified traffic volume and CO sampled predictors were subsequently defined in terms of an n-by-n geographic weights matrix, C, whose c ij values were 1 if the sampled i and j were deemed nearby and 0 otherwise. Adjusting this matrix by dividing each row entry by its row sum, with the row sums given by C1, converted this matrix-to-matrix W.

All residual estimates from the model were then evaluated in a spatial error (SE) model. An autoregressive model was employed that used a sampled variable, Y, as a function of nearby sampled Y values [that is, an autoregressive response (AR) or spatial linear (SL) specification] and/or the residuals of Y as a function of nearby Y residuals [that is, an AR or SE specification]. Euclidean distance between traffic volume and CO sampled geolocations was defined in terms of an *n*-by-*n* geographic weights matrix, C, whose c_{ij} values were 1 if the traffic volume and CO sampled, locations *i* and *j* were deemed nearby, and 0 otherwise. Adjusting this matrix by dividing each row entry by its row sum, with the row sums given by C1, converted this matrix to matrix W (Griffith, 2003).

Study design

In this study, the AADT values of Highways in Hillsborough County were identified, and the corresponding CO concentration values were obtained by ArcGIS Pro 2.9 software. Eigendecomposed autocorrelation was employed for both AADT and CO values to detect the autocorrelation distribution patterns. Stratification based on AADT value was applied, that is, high-AADT and median-AADT groups. Pearson correlation coefficients were detected with SAS software between AADT (including different strata of AADT) and CO concentration values to analyze the correlation. In addition, to provide disease-prevention information to susceptible populations, essential buildings within the 1

km buffer zones were geolocated in Google Maps and ArcGIS Pro 2.9. Possible ethical issues were avoided by obtaining open data from the governmental website.

RESULTS

The identified essential buildings include 62 educational facilities, 42 hospitals, healthcare centers and clinics and 11 senior centers. Among the educational facilities, there are one children's museum and 61 elementary, middle and high schools for children under 18 years of age. The map with all essential buildings was created as shown in Figure 3.

In the process of identifying roadways, 50 roadways with AADT values equal to or larger than 80,000 VPD were included (mean=142550, 95% CI of mean= [131455, 153646] VPD, SD=39041 VPD) (Table 1). In this study, 13 road sections had high AADT values (mean=193615 VPD, 95% CI of mean=[181273, 205958] VPD, SD=20425 VPD) (Table 1), and 37 sections had medium AADT values (mean=124608 VPD, 95%CI of mean=[116023, 133194] VPD, SD=25750 VPD) (Table 1). The length (mean= 2695.6 m, 95% CI of mean=[2046.5, 3344.8], SD=2284.1 m) of each roadway was used to determine the number of points when measuring CO concentration (Table 1). As for the length of sections for high AADT and medium AADT, the mean length was 1549.6 m (95%CI of mean=[939.2, 2160.0], SD=1010.1 m) for the high AADT and 3098.3 m (95%CI of mean=[2273.4, 3923.1], SD=2473.9 m) for medium AADT. The total number of points added on each section of the roadway was 424, with 363 points (min=1, max=34) on sections with AADT values between 80,000 VPD and 160,000 VPD and 61 points (min=1, max=11) on sections with AADT values above 160,000 VPD (Table 1) (Figure 2).

The AADT and CO concentration-related, stratified, r-determinant values were converted to a z-score. Z-score is a numerical measurement that describes a value's relationship to the mean of a group of values (Morán, 1950). The Z-score is measurable in terms of standard deviations from the mean. Researchers can employ a Z-score for investigating epidemiologic, scalable capture point, sentinel site, georeferenceable, AADT-related, regressively, prognosticatable model outputs along congested highways as it can indicate that the geosampled, estimator determinant's score is identical to the mean score in summary diagnostic forecasts (e.g., geolocations of aggregation/non-aggregation-oriented, county-level, hot/cold spots and their respective CO satellite synthesized covariates). The formula for calculating a z-score in an empirical, epidemiologic, scalable, AADT-related, county vulnerability-oriented, prognosticative model for optimally, heuristically, optimizing, targeting and prioritizing, georeferenceable, aggregation/non-aggregation-oriented, stratifiable, geosampled, CO-related, capture point, sentinel site,
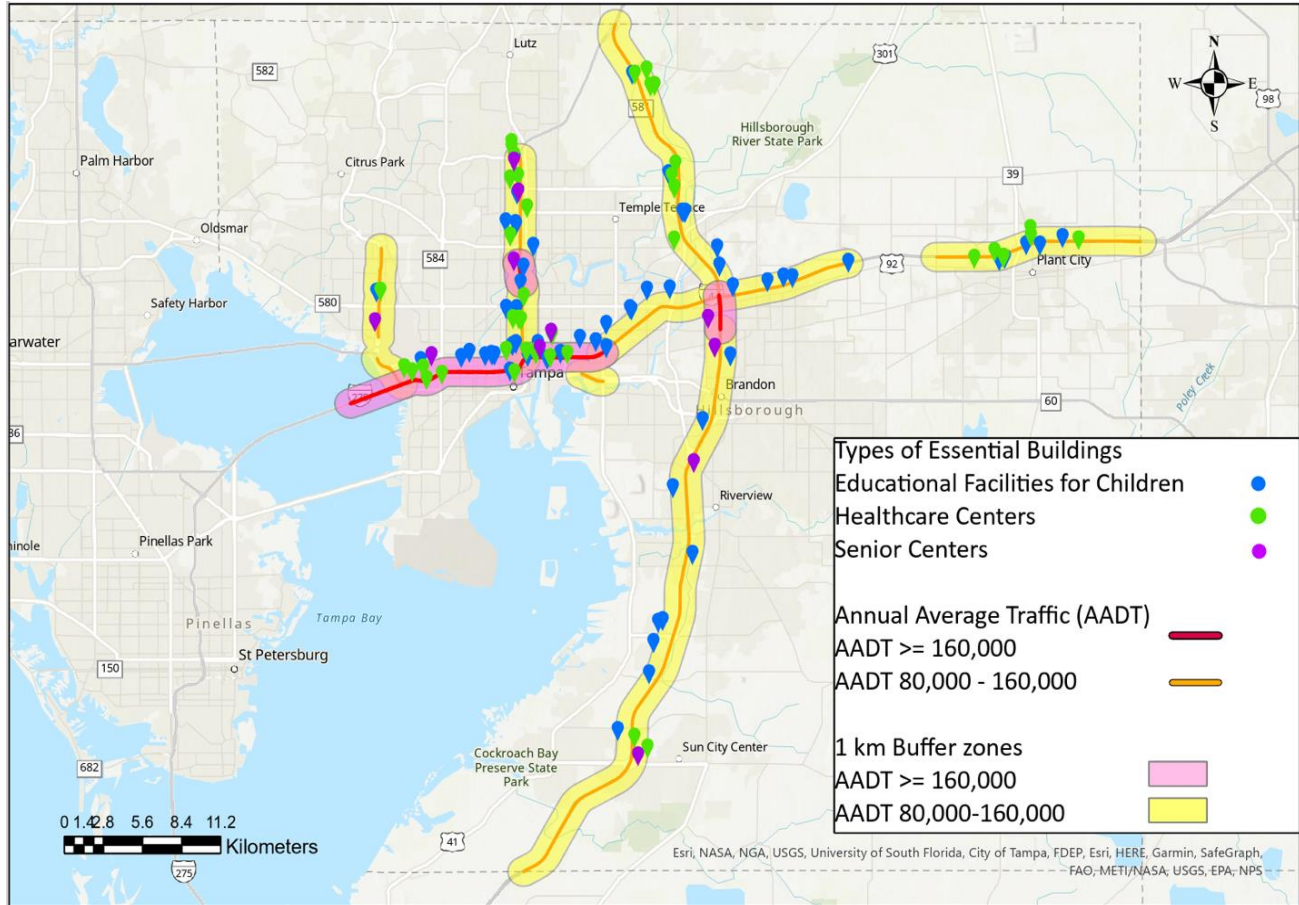
Figure 3. Map of identified essential buildings within the one-kilometer buffer of roadways in Hillsborough County, FL, 2023.
Source: Authors.

estimator determinants may be z = (x-μ)/σ, where x is the raw score, μ is the population mean, and σ is the population standard deviation in semiparametric, conditional, autoregressive, eigenvector eigen-geospace. As the formula reveals, defining the z-score in these analytical model estimator determinants may be ascertainable by any researcher or collaborator by calculating the raw score minus the population mean divided by the population standard deviation.

The autocorrelation results revealed a clustered distribution for AADT and CO concentration values. The Moran's I for the AADT was 0.956, with the expected index of -0.002, z-score of 22.709, and p-value < 0.001 (Figure 4). The z-score of 22.709 indicated a less than 1% likelihood that the clustered pattern of the AADT could be due to random chance. As for the CO concentrations, Moran's I was 0.973, with the expected index being -0.002, z-score of 23.064, and p-value < 0.001 (Figure 5). Given the z-score of 23.064, the possibility of the clustered pattern of the CO concentrations being the result of random chance could be less than 1%.

The n-by-1 vector $x = [x_1 \cdots x_n]^T$ contained

measurements of a quantitative variable for n spatial units and n-by-n spatial weighting matrix W. The formulation for Moran's I of spatial autocorrelation used in this research was:

$$I(x) = \frac{n \Sigma_{(2)} w_{ij} (x_i - \bar{x})(x_j - \bar{x})}{\Sigma_{(2)} w_{ij} \Sigma_{i=1}^{n} (x_i - \bar{x})^2}$$

where $\Sigma_{(2)} \sum_{i=1}^{n} \sum_{j=1}^{n}$ with $i \neq j$. The values w_{ij} were spatial weights stored in the symmetrical matrix W [that is, $(w_{ij} = w_{ji})$] that had a null diagonal ($w_{ii} = 0$). In this research, the matrix was initially generalized to an asymmetrical matrix W. Matrix W can be generalized by a non-symmetric matrix W^* by using $W = (W^* + W^{*T})/2$ (Griffith, 2003). Moran's I was rewritten using matrix notation:

$$I(x) = \frac{n}{1^t W 1} \frac{x^T HHWHHx}{x^T HHx} = \frac{n}{1^T W 1} \frac{x^T HWHx}{x^T Hx}$$

Table 1. Descriptive statistics of samples, one sample t-test results of AADT values, independent samples t-test results of length and number of points of roadways.

Variable	n (%)	Mean [95% CI[c]]	Min	Max	SD	t-statistic	df	p-value	η2
AADT (VPD[f])						25.82[d]	49[d]	<.0001[d]	-
High-AADT[a] (VPD[f])	13 (26)	193615 [181273, 205958]	165000	228000	20425				
Medium-AADT[b] (VPD[f])	37 (74)	124608 [116023, 133194]	80800	158000	25750				
Total (VPD[f])	50 (100)	142550 [131454, 153645]	80800	228000	39041				
Length						-3.14[e]	46.72[e]	0.003[e]	0.17[e]
High-AADT[a]	13 (26)	1549.62 [939.23, 2160.01]	402.63	3466.6	1010.1				
Medium-AADT[b]	37 (74)	3098.29 [2273.44, 3923.14]	379.87	10418	2473.9				
Total	50 (100)	2695.64 [2046.49, 3344.78]	379.87	10418	2284.1				
Number of points						-3.11[e]	46.96[e]	0.003[e]	0.17[e]
High-AADT[a]	61 (14.4)	5 [3, 7]	1	11	3				
Medium-AADT[b]	363 (85.6)	10 [7, 13]	1	34	8				
Total	424 (100)	8 [6, 11]	1	34	8				

[a]AADT ≥ 160,000. [b]AADT < 160,000 and ≥ 80,000. [c]Confidence interval. [d]Result of one sample t-test. [e]Result of independent samples t-test. [f]Vehicles per day.

where $H=(I-11^{T}/n)$ was an orthogonal projector verifying that $H=H^2$ (that is, H was independent). Features of matrix W for analyzing traffic volume and CO sampled data include that it is a stochastic matrix, expressing each observed value y_i as a function of the average of the location i's nearby data, and allows a single spatial autoregressive parameter, ρ, to have a maximum value of 1.

Subsequently, a SAR model specification was used to describe the autoregressive variance uncertainty estimates. A spatial filter (SF) model specification was also used to describe both Gaussian and Poisson random variables. The resulting SAR model specification took on the following form:

$$Y = \mu(1 - \rho)1 + \rho WY + \varepsilon \quad (2)$$

where μ was the scalar conditional mean of Y, and ε was an n-by-1 error vector whose elements

were statistically independent and identically distributed (iid) normally random variates. The spatial covariance matrix for Equation (1), using the sampled traffic volume and CO covariates, was $E[(Y - \mu I)' (Y - \mu I)] = \Sigma = [(I - \rho W')(I - \rho W)]^{-1}\sigma^2$, where $E(\bullet)$ denoted the calculus of expectations, I was the n-by-n identity matrix denoting the matrix transpose operation, and σ^2 was the error variance.

However, we assumed that when a mixture of positive and negative spatial autocorrelation is present in a traffic volume and CO frequency model, a more explicit representation of both effects leads to a more accurate interpretation of empirical results. Alternately, the excluded traffic volume and CO sampled values may be set to zero, although if this is done, then the mean and variance must be adjusted. In this research, two different spatial autoregressive parameters appeared in the spatial covariance matrix model specification, which for an SAR model specification

became:

$$\Sigma = [(I - <\rho>_{diag} W')(I - <\rho>_{diag} W)]^{-1}\sigma^2 \quad (3)$$

where the diagonal matrix of autoregressive parameters, $<\rho>_{diag}$, contained two sampled parameters: ρ_+ for those traffic volume and CO sampled pairs displaying positive spatial dependency, and ρ_- for those habitat pairs displaying negative spatial dependency. For example, by letting $\sigma^2 = 1$ and employing a 2-by-2 regular square tessellation,

$$\Sigma = \left[\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} - \begin{pmatrix} \rho_+ & 0 & 0 & 0 \\ 0 & \rho_+ & 0 & 0 \\ 0 & 0 & \rho_- & 0 \\ 0 & 0 & 0 & \rho_- \end{pmatrix} \begin{pmatrix} 0 & \frac{1}{2} & \frac{1}{2} & 0 \\ \frac{1}{2} & 0 & 0 & \frac{1}{2} \\ \frac{1}{2} & 0 & 0 & \frac{1}{2} \\ 0 & \frac{1}{2} & \frac{1}{2} & 0 \end{pmatrix} \right]^2$$
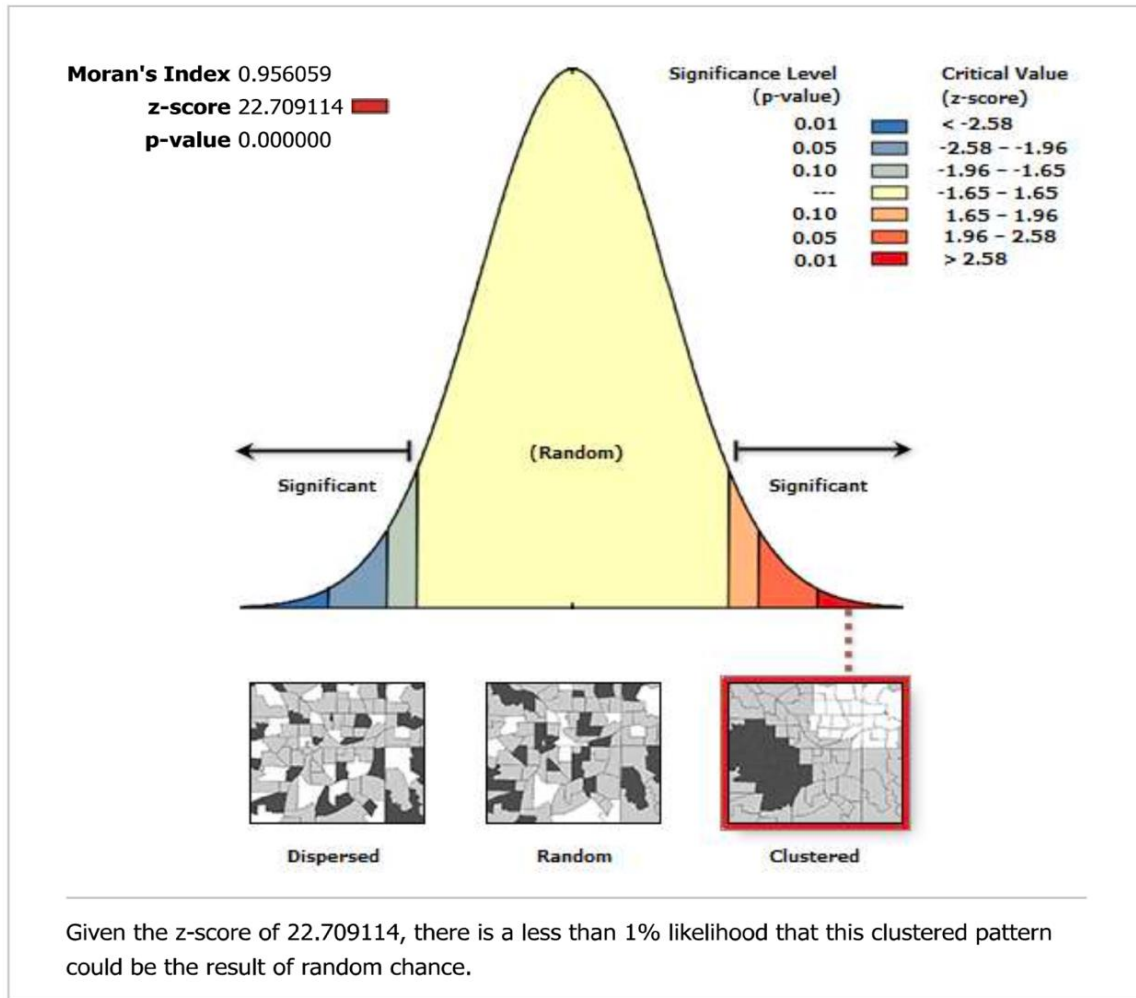
for the

Spatial Autocorrelation Report

Figure 4. Autocorrelation of AADT values.
Source: Authors.

vector $\begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{pmatrix}$, enabled positing a positive relationship between the traffic volume and CO sampled covariates y_1 and y_2, a negative relationship between covariates y_3 and y_4, and no relationship between covariates y_1 and y_3 and between y_2 and y_4. This covariance specification yielded:

$$Y = \mu(I - \rho_+ < I_+ >_{\text{diag}} - \rho_- < I_- >_{\text{diag}})\mathbf{1},$$
$$+(\rho_+ < I_+ >_{\text{diag}} + \rho_- < I_- >_{\text{diag}})WY + \varepsilon \quad (3)$$

where I_+ was a binary 0-1 indicator variable which denoted those traffic volume and CO sampled

determinants displaying positive spatial dependency, and I_- was a binary 0-1 indicator variable denoting those geolocations displaying negative spatial dependency, using $I_+ + I_- = 1$. Expressing the preceding 2-by-2 example in terms of Equation (2,3) yielded:

$$\begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{pmatrix} = \mu \left[\begin{pmatrix} 1&0&0&0 \\ 0&1&0&0 \\ 0&0&1&0 \\ 0&0&0&1 \end{pmatrix} - \rho_+ \begin{pmatrix} 1&0&0&0 \\ 0&1&0&0 \\ 0&0&0&0 \\ 0&0&0&0 \end{pmatrix} - \rho_- \begin{pmatrix} 0&0&0&0 \\ 0&0&0&0 \\ 0&0&1&0 \\ 0&0&0&1 \end{pmatrix} \right] \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} +$$

$$\left[\rho_+ \begin{pmatrix} 1&0&0&0 \\ 0&1&0&0 \\ 0&0&0&0 \\ 0&0&0&0 \end{pmatrix} + \rho_- \begin{pmatrix} 0&0&0&0 \\ 0&0&0&0 \\ 0&0&1&0 \\ 0&0&0&1 \end{pmatrix} \right] \begin{pmatrix} 0&\frac{1}{2}&\frac{1}{2}&0 \\ \frac{1}{2}&0&0&\frac{1}{2} \\ \frac{1}{2}&0&0&\frac{1}{2} \\ 0&\frac{1}{2}&\frac{1}{2}&0 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{pmatrix} + \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \varepsilon_3 \\ \varepsilon_4 \end{pmatrix}.$$

If either $\rho_+=0$ (and hence $I_+=0$ and $I_-=I$) or $\rho_- = 0$ (and

Spatial Autocorrelation Report

Moran's Index 0.973258
z-score 23.063833
p-value 0.000000

Significance Level (p-value)		Critical Value (z-score)
0.01		< -2.58
0.05		-2.58 – -1.96
0.10		-1.96 – -1.65
---		-1.65 – 1.65
0.10		1.65 – 1.96
0.05		1.96 – 2.58
0.01		> 2.58

(Random)

Significant   Significant

Dispersed   Random   Clustered

Given the z-score of 23.063833, there is a less than 1% likelihood that this clustered pattern could be the result of random chance.
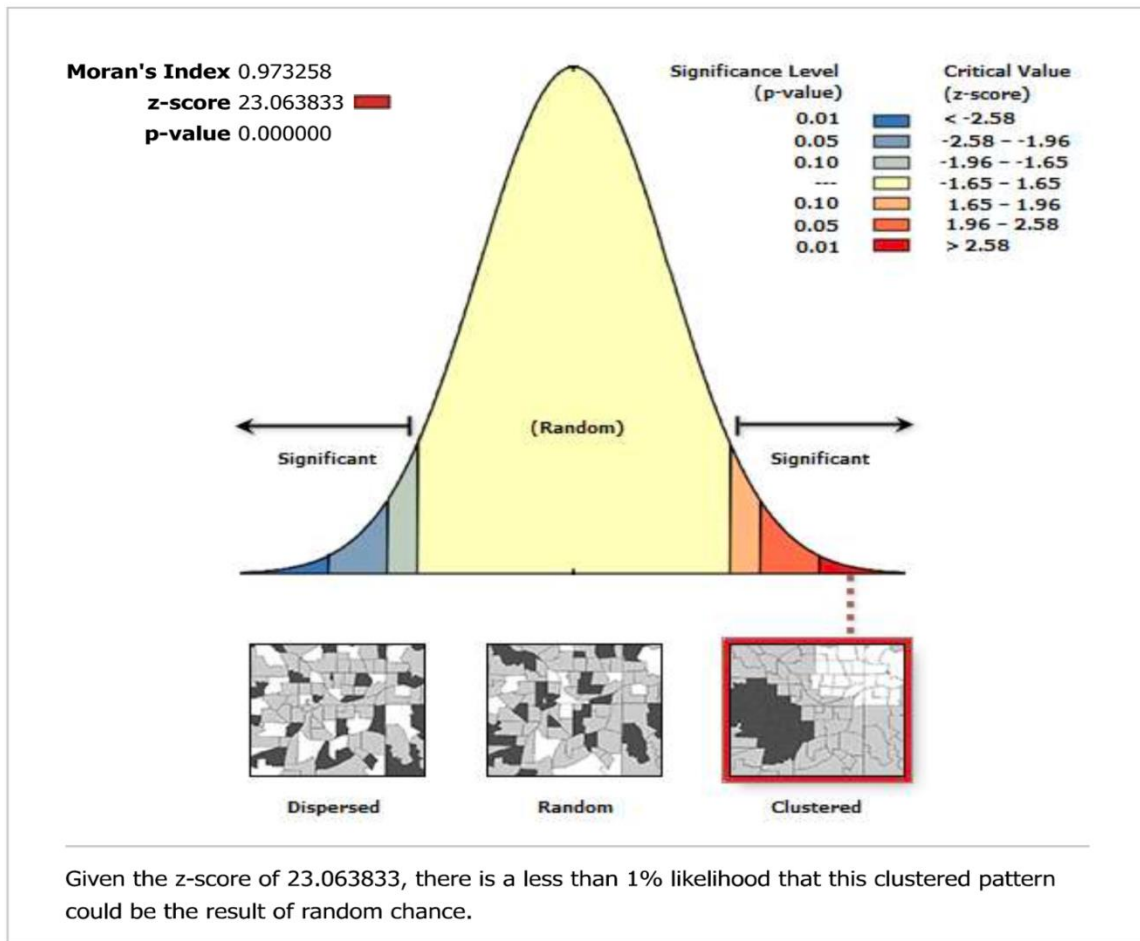
Figure 5. Autocorrelation of CO values.
Source: Authors.

hence $I_-=0$ and $I_+=I$), then Equation 4 is reduced to Equation 2. This indicator variables classification was made in accordance with the quadrants of the corresponding Moran scatterplot generated using the sampled traffic volume and CO sampled covariates sampled in the Hillsborough study site.

If positive and negative spatial autocorrelation processes counterbalance each other in a mixture, the sum of the two spatial autocorrelation parameters--$(\rho_+ + \rho_-)$ will be close to 0 (Griffith, 2003). In this research, Jacobian estimation was implemented by utilizing the differenced indicator traffic volume and CO sampled variables $(I_+ - \gamma I_-)$, estimating ρ_+ and γ with maximum likelihood techniques, and setting $\hat{\rho}_- = -\hat{\gamma}\hat{\rho}_+$. The Jacobian generalizes the gradient of a scalar-valued function of multiple variables, which itself generalizes the derivative of a scalar-valued function of a scalar (Griffith, 2003). A more complex traffic volume and CO sampled estimator determinant specification was then posited by

generalizing these binary indicator variables. We used $F: R^n \to R^m$ as, a function from Euclidean n-space to Euclidean m-space, which was generated using the distance between sampled traffic volume and CO specified covariates. Such a function was given by m covariate (that is, component functions), $y_1(x_1, xn)$, $y_m(x_1, xn)$. The partial derivatives of all these functions were organized in an m-by-n matrix, the Jacobian matrix J of F, which was as follows:

$$J = \begin{bmatrix} \dfrac{\partial y_1}{\partial x_1} & \cdots & \dfrac{\partial y_1}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \dfrac{\partial y_m}{\partial x_1} & \cdots & \dfrac{\partial y_m}{\partial x_n} \end{bmatrix}$$

This matrix was denoted by $J_F(x_1,..., x_n)$ and $\dfrac{\partial(y_1,...,y_m)}{\partial(x_1,...,x_n)}$. The ith row $(i=1,..., m)$ of this

matrix was the gradient of the ith component function y_i:(∇y_i). In this analysis, p was a sampled traffic volume and CO specified covariates in R^n and F (that is, sampled larval/pupal count) that were differentiable at p; its derivative was given by $J_F(p)$. The model described by $J_F(p)$) was the best linear approximation of F near the point p, in the sense that:

$$F(x) = F(p) + J_F(p)(x - p) + o(\|x - p\|) \qquad (4)$$

The spatial structuring was achieved by constructing a linear combination of a subset of the eigenvectors of a modified geographic weights matrix, using (I-11'/n) C (I-11'/n) that appeared in the numerator of the Moran's coefficient (MC) spatial autocorrelation can be indexed with an MC, a product-moment correlation coefficient (Griffith, 2003). A subset of eigenvectors was then selected with a stepwise regression procedure. Because (I - 11'/n) C (I - 11'/n)=E Λ E', where E is an *n*-by-*n* matrix of eigenvectors and Λ is an *n*-by-*n* diagonal matrix of the corresponding eigenvalues, the resulting traffic volume and CO sampled model specification was given by:

$$Y = \mu 1 + E_k \beta + \varepsilon, \qquad (5)$$

where *µ* the scalar mean of Y, Ek was an *n*-by-k matrix containing the subset of k $<<n$ eigenvectors selected with a stepwise regression technique, and β was a *k*-by-1 vector of regression coefficients.

A number of the eigenvectors were extractable from (I - 11'/n) C (I- 11'/n), which were affiliated with geographic patterns of the sampled traffic volume and CO-specified covariates in the study site, portraying a negligible degree of spatial autocorrelation. Consequently, only k of the *n* eigenvectors was of interest for generating a candidate set for a stepwise regression procedure. Candidate eigenvector represents a level of spatial autocorrelation that can account for the redundant information in orthogonal map traffic volume and CO sampled pattern. Of note is that because the 2-by-2 square tessellation rendered a repeated eigenvalue,

To identify georeferenced spatial clusters of traffic volume and CO-sampled geolocations, Thiessen polygon surface partitioning was generated in ArcGIS Pro to construct geographic neighbor matrices, which were also used in the spatial autocorrelation analysis. Entries in the matrix were 1 if two sampled traffic volumes and CO delineated geolocations shared a common Thiessen polygon boundary and 0, otherwise. Next, the linkage structure for each surface was edited to remove unlikely geographic neighbors to identify pairs of sampled variables sharing a common Thiessen polygon boundary. Attention was restricted to those map patterns associated with at least a minimum level of spatial autocorrelation, which, for implementation purposes, as defined by $|MC_j/MC_{max}| > 0.25$, where MC_j denoted the *j*th value

and MC_{max}, the maximum value of MC. This threshold value allowed two candidate sets of eigenvectors to be considered for substantial positive and substantial negative spatial autocorrelation, respectively. These statistics indicated that the detected negative spatial autocorrelation may be considered to be statistically significant based on a randomization perspective. Of note is that the ratio of the PRESS (that is, predicted error sum of squares) statistic to the sum of squared errors from the MC scatterplot trend line was well within two standard deviations of the average standard prediction error value for a traffic volume and CO sampled region in the Hillsborough study site.

The upper and lower bounds for a spatial matrix generated using Moran's *I* were given by $\lambda_{max}(n/1^T W 1)$ and $\lambda_{min}(n/1^T W 1)$, where λ_{max} and λ_{min} which were the extreme eigenvalues of $\Omega = HWH$. Hence, in this research, the eigenvectors of Ω were vectors with unit norm maximizing Moran's *I*. The eigenvalues of this matrix were equal to Moran's *I* coefficients of spatial autocorrelation post-multiplied by a constant. Eigenvectors associated with high positive (or negative) eigenvalues have high positive (or negative) autocorrelation (Griffith, 2003). The eigenvectors associated with eigenvalues with extremely small absolute values correspond to low residual geo-spatiotemporal autocorrelation, which was not suitable for defining spatial structures in the traffic volume and CO sampled data.

The diagonalization of the spatial weighting matrix generated from the field and remote-sampled traffic volume and CO sampled covariate coefficients consisted of finding the normalized vectors u_i, stored as columns in the matrix $U=[u_1 \cdots u_n]$, satisfying:

$$\Omega = HWH = U\Lambda U^T = \sum_{i=1}^{n} \lambda_i u_i u_i^T$$

where $\Lambda = diag(\lambda_1 \cdots \lambda_n)$, $u_i^T u_i = \|u_i\|^2 = 1$ and $u_i^T u_j = 0$ for $i \neq j$. Note that double centering of Ω implied that the eigenvectors u_i generated from the e sampled covariates were centered and at least one eigenvalue was equal to zero. Introducing these eigenvectors in the original formulation of Moran's *I* lead to:

$$I(x) = \frac{n}{1^T W 1} \frac{x^T HWHx}{x^T Hx} = \frac{n}{1^T W 1} \frac{x^T U\Lambda U^T x}{x^T Hx}$$

$$= \frac{n}{1^T W 1} \frac{\sum_{i=1}^{n} \lambda_i x^T u_i u_i^T x}{x^T Hx} \qquad (6)$$

Considering the centered vector $z = Hx$ and using the

properties of idempotence of *H*, Equation 6 was equivalent to:

$$I(x) = \frac{n}{1^T W 1} \frac{\sum_{i=1}^{n} \lambda_i z^T u_t u_i^T z}{z^T z} = \frac{n}{1^T W 1} \frac{\sum_{i=1}^{n} \lambda_i \|u_i^T z\|^2}{\|z\|^2}$$

(7)

As the eigenvectors u_i and the vector z were centered, Equation 7 was rewritten:

$$I(x) = \frac{n}{1^T W 1} \frac{\sum_{i=1}^{n} \lambda_i cor^2(u_i, z) \, var(z) n}{var(z) n}$$

$$= \frac{n}{1^T W 1} \sum_{i=1}^{n} \lambda_i cor^2(u_i, z)$$

(8)

In this research, *r* was the number of null eigenvalues of Ω ($r \geq 1$). These eigenvalues and corresponding eigenvectors were removed from Λ and *U* respectively. Equation 8 was then strictly equivalent to:

$$I(x) = \frac{n}{1^T W 1} \sum_{i=1}^{n-r} \lambda_i cor^2(u_i, z)$$

(9)

Moreover, it was demonstrated that Moran's *I* for a given eigenvector u_i was equal to $I(u_i) = (n/1^T W 1)\lambda_i$, so the equation was rewritten:

$$I(x) = \sum_{i=1}^{n-r} I(u_i) cor^2(u_i, z)$$

The term $cor^2(u_i, z)$ represented the part of the variance of *z* that was explained by u_i in the traffic volume and CO, estimator, determinant model $z = \beta_i u_i + e_i$. This quantity was equal to $\beta_i^2/n \, var(z)$. By definition, the eigenvectors u_i was orthogonal, and therefore, regression coefficients of the linear models $z = \beta_i u_i + e_i$ were those of the multiple regression model $z = U\beta + \varepsilon = \beta_i u_i + \cdots + \beta_{n-r} u_{n-r} + \varepsilon$.

The maximum value of *I* was obtained by all of the variations of *z*, as explained by the eigenvector u_1, which corresponded to the highest eigenvalue λ_1 in the spatial autocorrelation error matrix. In this research, $cor^2(u_i, z) = 1$ (and $cor^2(u_i, z) = 0$ for $i \neq 1$), and the maximum value of *I*, was deduced for Equation 9, which was equal to $I_{max} = \lambda_1(n/1^T W 1)$. The minimum value of *I* in the error matrix was obtained as all the variation of *z* was explained by the eigenvector u_{n-r} corresponding to the lowest eigenvalue λ_{n-r} generated in the traffic volume and CO frequency model. This minimum value was equal

to $I_{min} = \lambda_{n-r}(n/1^T W 1)$. If the sampled predictor variable was not spatialized, the part of the variance explained by each eigenvector was equal, on average, to $cor^2(u_i, z) = 1/n$-1. Because the field and remote-sampled traffic volume and CO variables in *z* were randomly permuted, it was assumed that we would obtain this result. The set of *n*! random permutations revealed that

$$E_R(I) = \frac{n}{1^T W 1 (n-1)} \sum_{i=1}^{n} \lambda_i = \frac{n}{1^T W 1 (n-1)} trace(\Omega)$$

. It was easily demonstrated that

$$trace(\Omega) = -\frac{1^T W 1}{n}$$

and it followed that

$$E_R(I) = -\frac{1}{n-1}$$

. The final model revealed a slight tendency for negative spatial autocorrelation in the traffic volume and CO sampled data.

A researcher or experimenter may construct space-time model AADT specifications based on Moran eigenvector space-time filters. Such an experiment can include zero/non-zero eigen-autocorrelation, which can refer to the correlation between existing CO at a capture point and other geosampled data (Cliff and Ord, 1973; Anselin, 1988; Griffith, 2003), for example, which can characterize data values that are not independent but rather are tied together in overlapping subsets within a given heavy trafficked highway geographic landscape. Such an experiment can summarize the various interpretations of autocorrelation with particular emphasis on its explanation as a prognosticatable, signature, seasonal, AADT map pattern. Eigenizable, latent, autocorrelation and AADT-related coefficients may also be determinable by employing the Moran Coefficient. Eigen-spatial filtering is a statistical method whose goal is to obtain enhanced and robust results in spatial data analysis by decomposing a spatial variable into the trend, a spatially structured random component (that is, a spatial stochastic signal), and random noise (Griffith, 2003). Additionally, an experimenter may also separate spatially structured random components from both trend and random noise, AADT modeling to sounder statistical inference and useful visualization of seasonal, topological land cover, and meteorological data. This separation procedure can involve eigenfunctions of the matrix version of the numerator of the Moran coefficient. Moran eigenvector spatial filtering conceptual materials may be presented using a computer code for implementing the procedure in SAS R or Python.

One sample t-test was conducted to analyze the AADT value distribution, and the result was significant, with a t-value (df=49) of 25.82 (p-value <.0001) (Table 1). Two independent samples t-tests were conducted to compare the length and number of points between two groups of roadways, which are high-AADT and medium-AADT groups. The equality of variances was tested with a F-value of 6 (p=0.002) (Table 1). There was a significant difference in lengths between the road sections with high AADT and medium AADT, with a t-value of -3.14 and a p-

value=0.003 (Table 1). The magnitude of the difference in the means (mean difference=-1548.67 m) was large with eta-squared=0.17 (Table 1). As for the difference in the number of points, the differences in the number of points on one section of the road between high AADT (mean=5, 95% CI of mean=[3, 7], SD=3) and medium AADT (mean=10, 95% CI of mean=[7, 13], SD=8) were significant with a t-value of -3.11 (p=0.003) (Table 1). The magnitude of the difference in the means (mean difference=-5) was large with eta-squared=0.17 (Table 1). The equality of variances was tested with an F-value of 6.21 (p=0.002) (Table 1).

In this experiment, spatial autocorrelation was defined as a particular relationship between the spatial proximity among observational AADT units and the numeric similarity among sampled CO values; positive spatial autocorrelation referred to situations in which the nearer the observational units (AADT and CO signatures), the more similar their values (and vice versa for its negative counterpart). The presence of non-zero spatial autocorrelation or dependence in our model meant that a certain amount of information was shared among neighboring geo-referenced traffic volume locations within our intervention site, and this feature violates the assumption of independent observations upon which many AADT standard statistical treatments are predicated. This latent autocovariance revolved around the nature and statistical significance of PCCs in the AADT model.

PCC was decomposed into two parts: Direct correlation (partial correlation) and indirect correlation (spatial cross-correlation) (Draper and Smith, 1981). The methodology was applied to determine the relationship between AADT and CO development so as to illustrate how to model this spatial cross-correlation phenomenon. This study is an introduction to developing spatial cross-correlation, and future geographical spatial analysis might benefit from forecast-oriented, AADT/CO signature models and vulnerability indexes.

The CO concentrations were measured by high-resolution daily Giovanni remote sensing data (Figure 6). The relationship between AADT value and CO concentrations was investigated by the PCC. Results showed a significant association with a p-value of 0.015 and a PCC (r) value of -0.119, n=424 (Table 2). This indicated a small effect of correlation between the two variables; when the annual average traffic volume increased, the annual average CO concentration decreased slightly. The same statistical test was applied to detect the relationship between high AADT and CO concentration values. There was a medium negative correlation between the two variables, r=-0.331, n=61, p-value=0.009, with higher AADT values indicating lower CO concentrations (Table 2). The correlation between medium AADT and CO concentration values was detected with the same method. The results showed a small positive correlation between the two variables, r =

0.142, n=363, p-value=0.007, with higher AADT values suggesting higher CO concentration (Table 2).

DISCUSSION

Initially, we identified 115 essential buildings located within one kilometer of distance from the medium- to high-traffic volumes in Hillsborough County, including elementary schools, middle schools, high schools, children's museums, healthcare service centers, clinics and senior centers. These buildings were pointed out as their visitor populations tend to be vulnerable to CO pollution (Raub, 1999; USEPA, 2015; Tang et al., 2019). The method of searching for potential CO pollution spots and correlating this to people's activities could be generalized to other geographic locations. We determined the one-kilometer buffer zone based on previous studies about the dissipation predictions and patterns of CO compound (Kamiński et al., 2007; USEPA, 2014). We believed that within a one-kilometer distance from the highways, the impacts of CO pollution could pose great threats to environmentally vulnerable populations, including children, senior people and patients (Raub, 1999; Tang et al., 2019; USEPA, 2014). Identifying buildings essential to the aforementioned populations was vital to this study as it provided information related to the vulnerability of health. This can be helpful in integrating the 10 Essential Public Health Services provided in the future in terms of accessing and monitoring environmental factors affecting the population (Centers for Disease Control and Prevention [CDC], 2023).

The Moran's *I* for AADT was 0.956 and for CO concentration was 0.973. The positive results of autocorrelation analyses revealed clustered patterns according to Moran's *I*, which indicated that the AADT and CO concentration values were clustered. One of the potential reasons for this is applying average traffic frequencies and CO concentration values. Employing mean or median values has the propensity to increase homogeneities in the parameters while reducing variances. In addition, as we selected the high-AADT and medium-AADT values as the independent variables, the roadways with lower AADT were not analyzed in this study. This further reduced heteroscedasticity in the AADT variable. In this case, the linear regression model may not be able to detect the association between the traffic volume and CO concentration values.

The PCC (r) between the independent and dependent variables was generated. The results of PCC for AADT and CO concentration was -0.119 (p-value=0.15); PCC r was -0.331 and 0.142 for high-AADT stratum and median-AADT (p < .05), respectively. The results revealed negative correlations between AADT and CO concentrations and also between the high-AADT group and corresponding CO concentrations, which are not
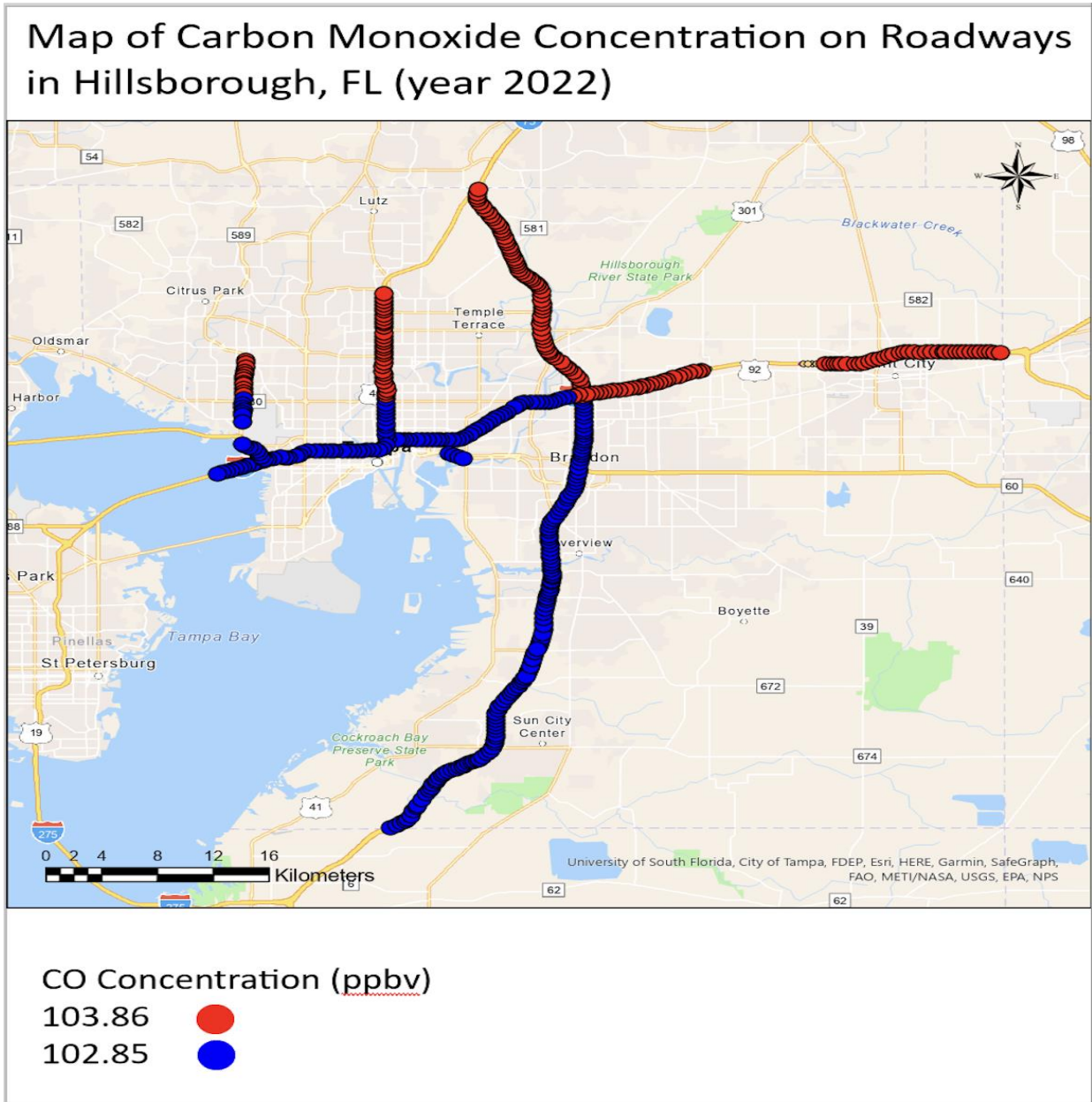
Figure 6. CO concentrations on measuring points on highways with medium and high AADT in Hillsborough County, Florida, 2022.
Source: Authors.

consistent with the hypothesis. These negative correlations can be explained by the clustered autocorrelations within AADT and CO concentrations, which biased the results of linear correlation. Additionally, as is shown in Figure 2, the highways with high AADT values tended to cluster in coastal areas, while Figure 6 illustrated that the CO concentrations were lower around the coastal areas. This can be explained by the weather and geographic conditions, including wind speed and directions, and the coastal location of Hillsborough

County, which also laid strong influences on the PCC (r). The coastal areas tend to have more winds than in-land areas as the winds are generated from movements of unevenly heated air above the ocean and earth, which is the situation in the downtown areas of Hillsborough County (United States Energy Information Administration [USEIA], 2022). The above factors can explain the biased correlation between the high AADT and CO concentrations and thus result in the negative association between overall AADT and CO concentration values.

Table 2. Descriptive statistics and correlations for study variables.

Variable	n	M	SD	p-value (r[c])					
				1	2	3	4	5	6
AADT (VPD[d])	424	134037	33178	-		-	-	-	-
CO	424	103.28	0.5	0.015 (-0.119[c])	-	-	-	-	-
High-AADT[a] (VPD[d])	61	189090	21586	-	-	-	-	-	
CO	61	102.93	0.28	-	-	0.009 (-0.33[c])	-	-	-
Medium-AADT[b] (VPD[d])	363	124785	24748	-	-	-	-	-	
CO	363	103.34	0.51	-	-	-	-	0.007 (0.142[c])	-

[a]AADT ≥ 160,000. [b]AADT < 160,000 and ≥ 80,000. [c]Pearson Correlation Coefficient (r). [d]Vehicles per day.

In the medium-AADT group, the values of AADT were positively correlated with CO concentrations, as we hypothesized. However, this does not mean that only the highways with medium levels of AADT are positively associated with CO pollution. As the weather conditions, such as wind speed and direction, vary among different areas in Hillsborough County, the correlations between the studied variables are biased by the weather conditions in different ways. Because the wind speeds tend to be higher along the coastal areas (USEIA, 2022), CO pollution decays faster than in the inland areas, which accounts for the spatial variances in CO concentrations throughout Hillsborough County. This indicates that involving weather parameters is necessary when analyzing the association between traffic conditions and related pollution. In addition, as extreme weather conditions are more common during Spring and Summer in Florida (FDOH, 2014; NWS, n.d.b), the winds and precipitations induced by them tend to increase the dissipation of CO compounds unevenly in terms of temporal variances (Njoku et al., 2022; NWS, n.d.a; Pan et al., 2016).

The independent t-tests comparing the two AADT groups revealed significantly different means of length and number of points in high-AADT and medium-AADT groups. In this study, the roadways with high AADT were shorter in length and involved fewer measuring points. Also, the standard deviation of length and number of points of the high-AADT group were smaller than those of the medium-AADT group. Therefore, heterogeneities in distribution patterns may also alter the association between AADT and CO concentration values. The aforementioned influential elements, including weather conditions, geographical factors, and sampling differences, altered the spatial-temporal variances and biased the results of the correlation between the independent and dependent variables in this study.

To reduce CO concentrations around essential buildings, a better building design that promotes ventilation in the building can effectively reduce exposure to CO pollution when CO compounds aggregate inside the building (USEPA, 2015). However, during rush hours, windows and doors to the outside environment should be closed, while using the heating, ventilating, and air conditioning (HVAC) system with high minimum efficiency reporting value (MERV) rating filters should be equipped by all the essential buildings (USEPA, 2015). By altering wind direction and speed in downwind areas, planting full coverage vegetation alongside the traffic roads can effectively reduce air pollution (Deshmukh et al., 2019; USEPA, 2015). Moreover, the vegetation buffers should cover from the top of the canopies to the ground to provide the best filtration result in reducing air pollutants such as CO compounds (Deshmukh et al., 2019; USEPA, 2015). Additionally, constructing new buildings farther than at least 500 feet away from the main traffic roads was also recommended to avoid exposure to traffic-induced pollution by some state governments (USEPA, 2015). In the long term, encouraging public transportation can decrease the number of private vehicles on the road by increasing the efficiency of transportation and thus alleviate CO pollution (Department of Ecology State of Washington [DESW], n.d.).

We noticed that there are some potential biases and limitations in this study. The first bias might be the misclassification bias. This is one of the common biases when doing health research (Althubaiti, 2016). This study might be subjected to this bias because the CO concentration measuring points were added manually from ArcGIS Pro. Since the process might not be precise, the measuring points of CO concentration values might not occur on the same coordinates

of the traffic volumes. The second bias is the clustered independent and dependent variables. As the traffic volume and CO concentration values obtained were clustered due to the calculation method of their original datasets, the correlation coefficient between them might be biased due to non-homoscedasticity. Additionally, potential effect modifiers that might impact the results in this study include wind speed and direction, humidity, temperature and precipitations (Guo et al., 2021; Njoku et al., 2022; NWS, n.d.a; Pan et al., 2016; USEPA, 2014). These parameters, especially wind speed may bias the results due to their influences on CO dissipation and decay rate (Njoku et al., 2022; NWS, n.d.a; Pan et al., 2016; USEPA, 2014). Moreover, local road traffic-related CO pollution was not analyzed in this study, although this can cause severe health problems for the people living nearby (Razavi-termeh et al., 2019).

In future research efforts, we would like to implement a signature spatial-temporal interpolation model stratified by land use land cover (LULC) and elevation. The study will create a one-kilometer squared vulnerability stratified grid. The grid will be stratified by levels of CO concentration and traffic volume throughout Orange County, LA, US, and Beijing, China. This method has the propensity to not only mitigate the bias due to non-homoscedasticity within interested variables but also complete the coverage of traffic volume measurement. Also, weather conditions should be included to minimize the potential bias caused by meteorological parameters in future studies.

Conclusion

In conclusion, the study aimed to detect autocorrelation distribution patterns and the association of AADT and daytime CO in Hillsborough County, FL. In this study, eigenfunction eigendecomposition algorithms were applied to detect the autocorrelation within the AADT and daytime CO concentration variables. The results of Moran's *I* were positive, indicating clustered patterns and propensity of heteroscedasticity existing in both AADT and CO concentration variables, largely due to the homogenization resulting from the averaging process of the original datasets. PCC was employed to analyze the association between AADT and CO concentrations, and the results showed a negative correlation between AADT and CO concentration values. However, after stratification, the medium-AADT stratum was positively correlated with CO concentration values. Considering the PPC (r) between the high-AADT stratum and corresponding CO concentrations was negative, this may bias the association between AADT and CO concentrations. It is possible that the results of this study may have been affected by unadjusted meteorological factors. Despite this, the statistical methods of testing distribution patterns and the association between

variables have been demonstrated in this study with optimal outcomes. A signature spatial-temporal interpolation model could be established in future research to involve climatic parameters to detect geo-spatiotemporal errors and reveal the correlation between traffic activities and air pollution by improving sampling techniques and analyzing meteorology modification.

CONFLICT OF INTERESTS

The authors have not declared any conflict of interest.

REFERENCES

Abedian S, Mirsanjari MM, Salmanmahiny A (2021). Investigating the effect of suburban buses on traffic flow and carbon monoxide emission by Aimsun simulation software. Journal of the Indian Society of Remote Sensing 49(6):1319-1330. https://doi.org/10.1007/s12524-020-01289-z

Akhtar M, Moridpour S (2021). A review of traffic congestion prediction using Artificial Intelligence. Journal of Advanced Transportation pp. 1-18. https://doi.org/10.1155/2021/8878011

Ali G, Sajjad M, Kanwal S, Xiao T, Khalid S, Shoaib F, Gul HN (2021). Spatial-temporal characterization of rainfall in Pakistan during the past half-century (1961-2020). Scientific Reports 11(1):6935. https://doi.org/10.1038/s41598-021-86412-x

Althubaiti A (2016). Information bias in health research: definition, pitfalls, and adjustment methods. Journal of Multidisciplinary Healthcare, 9:211-217. https://doi.org/10.2147/JMDH.S104807

Anselin L (1988). A test for spatial autocorrelation in seemingly unrelated regressions. Economics Letters 28(4):335-341. https://doi.org/10.1016/0165-1765 (88)90009-2

Benjamin J (1986). A time-series forecast of average daily traffic volume. Transportation Research Part A: General 20(1):51-60. https://doi.org/10.1016/0191-2607 (86)90015-4

Biswal A, Singh V, Malik L, Tiwari G, Ravindra K, Mor S (2023). Spatially resolved hourly traffic emission over megacity Delhi using advanced traffic flow data. Earth System Science Data 15(2):661-680. https://doi.org/10.5194/essd-15-661-2023

Centers for Disease Control and Prevention (2023). 10 Essential public health services. https://www.cdc.gov/publichealthgateway/publichealthservices/essentialhealthservices.html

Cliff AD, Ord JK (1973). Spatial autocorrelation. London: Pion.

Cressie NAC (1993). Statistics for spatial data revised edition. New York: John Wiley & Sons, Inc.

Datagov (2022). Annual average daily traffic (AADT): Beginning 1977. https://catalog.data.gov/dataset/annual-average-daily-traffic-aadt-beginning-1977

Department of Ecology State of Washington (n.d.). Reducing air pollution from cars. https://ecology.wa.gov/Issues-and-local-projects/Education-training/What-you-can-do/Reducing-car-pollution#:~:text=Reducing%20the%20amount%20of%20miles,riding%20the%20bus%20or%20train.

Deshmukh P, Isakov V, Venkatram A, Yang B, Zhang KM, Logan R, Baldauf R (2019). The effects of roadside vegetation characteristics on local, near-road air quality. Air Quality, Atmosphere and Health, 12(3):259-270. https://doi.org/10.1007/s11869-018-0651-8

Draper NR, Smith H (1981). Applied Regression analysis. John Wiley & Sons.

Federal Highway Administration (2018). Traffic data computation method. https://www.fhwa.dot.gov/policyinformation/pubs/pl18027_traffic_data_pocket_guide.pdf

Florida Climate Center (FCC) (2020). Average wind speed. https://climatecenter.fsu.edu/products-services/data/other-normals/average-wind-speed

Florida Department of Health (FDOH) (2014). Climate and health. https://www.floridahealth.gov/environmental-health/climate-and-health/index.html

Florida Department of Highway Safety and Motor Vehicles (2022). Registered vehicles and vessels, January 2022. https://www.flhsmv.gov/pdf/vehicle-vesselreports/cvr_01_2022_ada.pdf

Florida Department of Transportation (2022). Annual average daily traffic [Data set]. Retrieved July 5, 2023, at https://ftp.fdot.gov/file/d/FTP/FDOT/co/planning/transtat/gis/TRANSTAT_metadata/aadt.shp.xm

Freedman DR, Pisani R, Purves R (2007). Statistics fourth edition, William Warder Norton & Company, New York.

Griffith DA (2003). Spatial autocorrelation and spatial filtering gaining understanding through theory and scientific visualization first edition 2003. Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-540-24806-4

Guo H, Li X, Li W, Wu J, Wang S, Wei J (2021). Climatic modification effects on the association between PM1 and lung cancer incidence in China. BioMed Central Public Health 21(1):880-880. https://doi.org/10.1186/s12889-021-10912-8

Hauptman M, Gaffin JM, Petty CR, Sheehan WJ, Lai PS, Coull B, Gold DR, Phipatanakul W (2020). Proximity to major roadways and asthma symptoms in the School Inner-City Asthma Study. The Journal of Allergy and Clinical Immunology 145(1):119-126.e4. https://doi.org/10.1016/j.jaci.2019.08.038

Ierodiakonou D, Coull BA, Zanobetti A, Postma DS, Boezen HM, Vonk JM, McKone EF, Schildcrout JS, Koppelman GH, Croteau-Chonka DC, Lumley T, Koutrakis P, Schwartz J, Gold DR, Weiss ST (2019). Pathway analysis of a genome-wide gene by air pollution interaction study in asthmatic children. Journal of Exposure Science & Environmental Epidemiology 29(4):539-547. https://doi.org/10.1038/s41370-019-0136-3

Jacob BG, Krapp F, Ponce M, Zhang N, Caliskan S, Hasemann J, Griffith DA, Gotuzzo E, Novak RJ (2013). A Bayesian Poisson specification with a conditionally autoregressive prior and a residual Moran's coefficient minimization criterion for quantitating leptokurtic distributions in regression-based multi-drug resistant tuberculosis treatment protocols. Journal of Public Health and Epidemiology 5(3):122-143.

Kamiński W, Kucharski M, Tomczak E, Petera J (2007). Modelling Of carbon monoxide dispersion along roads with the use of the finite element method. In WIT Transactions on Ecology and the Environment P 101. https://doi.org/10.2495/AIR070071

Kingsley S L, Eliot MN, Carlson L, Finn J, MacIntosh DL, Suh HH, Wellenius GA (2014). Proximity of US schools to major roadways: a nationwide assessment. Journal of Exposure Science & Environmental Epidemiology 24(3):253-259. https://doi.org/10.1038/jes.2014.5

Le NH, Ly BT, Thai PK, Pham GH, Ngo IH, Do VN, Le TT, Nhu LV, Son HD, Nguyen YLT, Pham DH, Vu TV (2021). Assessing the impact of traffic emissions on fine particulate matter and carbon monoxide levels in Hanoi through COVID-19 social distancing periods. Aerosol and Air Quality Research 21(10):210081. https://doi.org/10.4209/aaqr.210081

Meyer D (2022). Time averaged map of carbon monoxide, mole fraction in air (daytime/ascending, airs-only) daily 1 deg. @1000hPa [airs airs3std v006] ppbv over 2022-01-01 - 2022-12-31 [Data set]. NASA/GSFC, Greenbelt, MD, USA, NASA Goddard Earth Sciences Data and Information Services Center (GES DISC). Retrieved July 9 2023 at https://giovanni.gsfc.nasa.gov/giovanni/#service= TmAvMp& starttime=2022-01-01T00:00:00Z&endtime=2022-12-31T23:59: 59Z&data=AIRS3STD_006_CO_VMR_A(z%3D1000)&dataKeyword= CO

Morán PAP (1950). Notes on continuous stochastic phenomena. Biometrika 37(1-2):17-23. https://doi.org/10.1093/biomet/37.1-2.17

National Aeronautics and Space Administration (n.d.). Giovanni measurement definitions: Carbon monoxide. https://disc.gsfc.nasa.gov/information/glossary?keywords=giovanni% 20measurements&title=Giovanni%20Measurement%20Definitions:% 20Carbon%20Monoxide

National Weather Service (n.d.a). Clearing the air on weather and air

quality. https://www.weather.gov/wrn/summer-article-clearing-the-air

National Weather Service (n.d.b). Florida thunderstorm season. https://www.weather.gov/tbw/RainySeason#:~:text=West%20Central %20and%20Southwest%20Florida,Zone%20just%20to%20the%20so uth.

Njoku EA, Akpan PE, Effiong AE, Babatunde IO, Owoseni OA, Olanrewaju JO (2022). Evaluation of geostatistical and multiple regression models for assessment of spatial characteristics of carbon monoxide concentration in a data-limited environment. Applied Geography (Sevenoaks), 149:102816. https://doi.org/10.1016/j.apgeog.2022.102816

National Centers for Environmental Information (NOAA) (2023). Climate at a glance: county time series. National Oceanic and Atmospheric Administration. https://www.ncei.noaa.gov/access/monitoring/climate-at-a-glance/county/time-series/FL-057/tavg/12/6/2022-2022

Orth S, Russell AG (2023). Assessment of light-duty versus heavy-duty diesel on-road mobile source emissions using general additive models applied to traffic volume and air quality data and COVID-19 responses. Journal of the Air & Waste Management Association, 73(5):374-393. https://doi.org/10.1080/10962247.2023.2185315

Pan L, Yao E, Yang Y (2016). Impact analysis of traffic-related air pollution based on real-time traffic and basic meteorological information. Journal of Environmental Management 183(Pt3):510-520. https://doi.org/10.1016/j.jenvman.2016.09.010

Prunicki M, Stell L, Dinakarpandian D, de Planell-Saguer M, Lucas RW, Hammond SK, Balmes JR, Zhou X, Paglino T, Sabatti C, Miller RL, Nadeau KC (2018). Exposure to NO$_2$, CO, and PM$_{2.5}$ is linked to regional DNA methylation differences in asthma. Clinical Epigenetics, 10:2. https://doi.org/10.1186/s13148-017-0433-4

Raub J (1999). Carbon monoxide second edition. World Health Organization. https://apps.who.int/iris/bitstream/handle/10665/42180/WHO_EHC_2 13.pdf

Razavi-termeh SV, Sadeghi-niaraki A, Choi SM (2021). Spatial modeling of asthma-prone areas using remote sensing and ensemble machine learning algorithms. Remote Sensing (Basel, Switzerland), 13(16):3222. https://doi.org/10.3390/rs13163222

Tang CS, Chuang KJ, Chang TY, Chuang HC, Chen LH, Lung SCC, Chang LT (2019). Effects of personal exposures to micro-and nano-particulate matter, black carbon, particle-bound polycyclic aromatic hydrocarbons, and carbon monoxide on heart rate variability in a panel of healthy older subjects. International Journal of Environmental Research and Public Health 16(23):4672. https://doi.org/10.3390/ijerph16234672

Tarigan APM, Suryati I, Gusrianti D (2018). A spatial analysis of the dispersion of transportation induced carbon monoxide using the Gaussian line source method. IOP Conference Series: Earth and Environmental Science 126(1):12120. https://doi.org/10.1088/1755-1315/126/1/012120

United States Census Bureau (2021). Hillsborough County, Florida. https://data.census.gov/profile/Hillsborough_County,_Florida?g=050X X00US12057

United States Consumer Product Safety Commission (USCPSC) (n.d.). Carbon monoxide questions and answers. https://www.cpsc.gov/Safety-Education/Safety-Education-Centers/ Carbon-Monoxide-Information-Center/Carbon-Monoxide-Questions-and-Answers#:~:text=As%20CO%20levels%20increase%20and, unconsciousness%2C%20and%20death%20are%20possible

United States Energy Information Administration (USEIA) (2022). Wind explained. https://www.eia.gov/energyexplained/wind/#:~:text= Project%20(public%20domain),The%20daily%20wind%20cycle,than %20it%20does%20over%20water.

United States Environmental Protection Agency (USEPA) (2014). Near roadway air pollution and health: Frequently asked questions. https://www.epa.gov/sites/default/files/2015-11/documents/ 420f14044_0.pdf

United States Environmental Protection Agency (USEPA) (2015). Best practices for reducing near-road pollution exposure at schools. https://www.epa.gov/sites/default/files/2015-10/documents/ochp_2015_near_road_pollution_booklet_v16_508.pdf

United States Environmental Protection Agency (USEPA) (2022). Basic

information about carbon monoxide (CO) outdoor air pollution. https://www.epa.gov/co-pollution/basic-information-about-carbon-monoxide-co-outdoor-air-pollution

United States Environmental Protection Agency (USEPA) (2023a). Overview of carbon monoxide (CO) air quality in the United States. https://www.epa.gov/system/files/documents/2022-08/CO_2021.pdf

United States Environmental Protection Agency (USEPA) (2023b). The links between air pollution and childhood asthma. https://www.epa.gov/sciencematters/links-between-air-pollution-and-childhood-asthma

Yu S, Yan Z, Che N, Zhang X, Ge R (2015). Impact of carbon monoxide/heme oxygenase on hydrogen sulfide/cystathionine-γ-lyase pathway in the pathogenesis of allergic rhinitis in guinea pigs. Otolaryngology-Head and Neck Surgery 152(3):470-476. https://doi.org/10.1177/0194599814567112

Voiland A (2015). Fourteen years of carbon monoxide from MOPITT. https://climate.nasa.gov/news/2291/fourteen-years-of-carbon-monoxide-from-mopitt/

Wang S, Ma Y, Wang Z, Wang L, Chi X, Ding A, Yao M, Li Y, Li Q, Wu M, Zhang L, Xiao Y, Zhang Y (2021). Mobile monitoring of urban air quality at high spatial resolution by low-cost sensors: Impacts of COVID-19 pandemic lockdown. Atmospheric Chemistry and Physics 21(9):7199-7215. https://doi.org/10.5194/acp-21-7199-2021

Wen D, Zhai W, Xiang S, Hu Z, Wei T, Noll, KE (2017). Near-roadway monitoring of vehicle emissions as a function of mode of operation for light-duty vehicles. Journal of the Air & Waste Management Association 67(11):1229-1239. https://doi.org/10.1080/10962247.2017.1330713