GIS-BASED REGRESSION ANALYSIS OF AIR QUALITY AND LAND SURFACE TEMPERATURE IN METRO MANILA AIRSHEDS

M. J. Broqueza[1], L. M. Cruz[1*], M. Mana-ay[1], A. M. Jamboy[1], R. V. Ramos [1,2]

[1] Department of Geodetic Engineering, University of the Philippines, Diliman, Quezon City, Philippines - (mjbroqueza, ldcruz6, mrmanaay, ajjamboy)@up.edu.ph

[2] Training Center for Applied Geodesy and Photogrammetry, University of the Philippines, Diliman, Quezon City, Philippines - rvramos@up.edu.ph

KEY WORDS: Land Surface Temperature, Air Quality, MODIS, Sentinel-5P, Ordinary Least Squares

ABSTRACT:

This study examines the spatial relationship of land surface temperature (LST) derived from MODIS, elevation, and air quality parameters (CO, NO_2, and SO_2) derived from Sentinel 5P in the Metro Manila airsheds from January 2019 to March 2023. Using Ordinary Least Squares (OLS) and Generalized Linear Regression (GLR), mean LST and heat index from the Philippine Atmospheric, Geophysical, and Astronomical Services Administration (PAGASA) ground stations exhibit a strong positive correlation, allowing the use of LST for further analysis. Across different combinations between LST, elevation, and air quality parameters, a weak to low negative correlation was seen between DEM to LST, CO, and NO_2. In addition, weak to low positive correlation was seen between LST to CO and NO_2. Almost no correlation was found between DEM and SO_2, and LST and SO_2. These results may be unreliable due to overfitting, non-stationarity, bias, or misspecification as implied by their statistical parameters. To enhance the reliability, it is recommended to investigate additional air quality parameters such as Normalized Difference Built-Up Index (NDBI) as high LST, CO, and NO_2 have shown clustering in urban areas of Metro Manila. Moreover, it suggests exploring other regression modeling and methodologies, such as training and test sets, to identify the best-fit model. In conclusion, this study provides an exploratory foundation for future research and comparative assessment on using different methods for modeling these variables. This comprehensive approach enhances understanding of the complex interplay between temperature, elevation, and air quality, aiding the development of informed urban climate adaptation strategies.

1. INTRODUCTION

1.1 Airsheds in the Philippines

Airsheds are defined as geographical areas within which the movement and quality of air are relatively uniform. Airsheds are crucial in understanding air quality, atmospheric conditions, and the dispersion of pollutants, which are essential for environmental management and planning (EMB, 2019).

Airsheds in the Philippines were delineated by different laws such as the Department of Environment and Natural Resources (DENR) Administrative Order (DAO), Memorandum Circular (MC) and Special Orders (SO) based on the following factors: geographical boundaries, meteorological factors, topography and physical features, and pollution sources.

Once an airshed is defined, it serves as the basis for the implementation of air quality management strategies and measures. The DENR and its Environmental Management Bureau (EMB), in collaboration with local government units and other stakeholders, develop airshed management plans and regulations to monitor, control, and reduce air pollution within the designated airshed (EMB, 2019).

In the Philippines, air quality management is primarily governed by the Clean Air Act of 1999 (Republic Act No. 8749), which aims to protect and preserve the country's air resources. The DENR and EMB are responsible for implementing the provisions of the law, including the establishment of airsheds.

1.2 Land Surface Temperature and Air Quality

Fuladlu and Altan in 2021, with the use of remote sensing, investigated the relationship between LST, air pollutants, and air quality in Tehran over a year (January to December 2020). Findings revealed that air pollutant concentrations are highest in the Tehran metropolis and core urban area. A negative correlation between the $PM_{2.5}$, SO_2, NO_2, and DEM was highlighted and increasing the DEM also negatively affects the concentration of LST, CO, and O_3 values.

Another study (Wang, Guo, & Han, 2021) examines the relationship between land surface temperature (LST) and air quality, focusing on the spatiotemporal patterns of Urban Heat Island Intensity (UHII) and six key air pollutants (CO, NO_2, O_3, $PM_{2.5}$ PM_{10}, SO_2) spanning 2015 to 2019. Among the considered factors, LST has the most influence, followed by vegetation cover, geographical location, elevation, and economic development intensity. This research underscores the complex relationship between urban heat islands, air quality, and various contributing factors across different regions and time frames. The results show that Urban Heat Island Intensity (UHII) and air pollutant concentrations are influenced by both natural features and socio-economic development. Higher LSTs and elevation contribute to increased daytime UHII, while urban expansion, higher per capita Gross Domestic Product (GDP), and greater Normalized Difference Vegetation Index (NDVI) lead to heightened nighttime UHII. Urban areas with higher population density and per capita GDP exhibit greater NO_2 pollution, while proximity to the ocean, enhanced vegetation, and elevated terrain correspond to lower air pollutant levels. (Wang et al, 2021).

* Corresponding author

https://doi.org/10.5194/isprs-archives-XLVIII-4-W8-2023-77-2024

1.3 Research Objectives and Significance

This study aims to determine the relationship between land surface temperature and air quality indicators, concentrations of CO, NO_2, and SO_2 pollutants. Specifically, it aims to determine the statistical significance of the correlation between Heat Index (HI) and land surface temperature (LST). This also seeks to determine whether there is a significant relationship between LST and selected air quality indicators. Additionally, it assesses and quantifies the impact of the Digital Elevation Model (DEM) on the outcomes and findings of this conducted research. Lastly, it comprehensively integrates and analyzes diverse statistical data, encompassing comparisons, regression, and correlations of LST, air quality, and elevation of Metro Manila airsheds (shown in Figure 1) from 2019-2023.

Understanding the relationship of land surface temperature, digital elevation model, and air quality helps to provide findings that can contribute to climate change adaptation strategies, particularly in urban areas susceptible to extreme heat and poor air quality. Recognizing patterns and connections between these factors aids in devising measures to ease urban heat island impact and manage air quality.

2. MATERIALS AND METHODS

2.1 Datasets and Pre-Processing

The heat index was provided by the Philippine Atmospheric, Geophysical and Astronomical Services Administration (PAGASA). This data encompassed the highest and average temperatures (in degree Celsius) recorded daily from the years 2019 to 2023 in tabulated format. These temperature readings were gathered from various ground stations situated in strategic locations, including NCR (specifically Quezon City, Pasay, and Manila), as well as areas in Pampanga, Bataan, Cavite, and Rizal.

Land Surface Temperature (LST) was derived from MODIS satellite images, retrieved daily with 1x1km spatial resolution (Wan, MODIS WE. The unit of measurement of this data is in °C. Consequently, air pollutants such as Carbon Monoxide (CO), Nitrogen Dioxide (NO_2), and Sulfur Dioxide (SO_2) were derived from Sentinel-5 Precursor TROPOMI satellite images, retrieved daily with 7x7 km and 3.5x 7 km spatial resolution for CO, and NO2 and SO2, respectively. These air pollutants have a mol/m² unit of measurement. In addition, the Digital Elevation Model (DEM) was retrieved using an image derived from Interferometric Synthetic Aperture Radar (IFSAR), originally having a 10x10 m spatial resolution.

The datasets were reprojected into the appropriate reference system for Metro Manila airsheds, WGS84 UTM Zone 51N, resample to a 1x1 km spatial resolution, and clipped to the extent of the airsheds. Daily datasets were aggregated to monthly mean values to reduce the gaps in the datasets due to clouds and data capture errors. These raster datasets were then converted into point vector files and spatially joined with each other.

The study's specific time frame spans from January 2019 to March 2023, chosen to align with data availability. This period encompasses Land Surface Temperature (LST) data until March 2023 and Heat Index (HI) data from January 2019. The timeframe also captures notable events like the onset of the COVID-19 pandemic (e.g., March 2020 to present), enabling examination of air quality variations during different pandemic phases. By comparing air quality data across years, the study seeks to identify potential shifts in pollution levels. Satellite imagery for 2018 is limited to December, hampering its utility for regression analysis.

The spatial scope of the study primarily revolves around the central part of Luzon, which is among the three major islands composing the Philippines. According to DENR Administrative Order 2011-11, the *"Metro Manila Airshed"* has been divided into three distinct airsheds: the National Capital Region (NCR) airshed, the Cavite-Laguna Rizal airshed, and the Bulacan-Pampanga-Bataan airshed. Metro Manila and several industrialized regions in the Philippines are renowned for having substantial pollution levels. The convergence of emissions from vehicles, industrial operations, and insufficient air quality control can result in heightened air pollution levels in these urban zones.

2.2 Processing of LST and Air Quality Parameters

The accuracy of Land Surface Temperature (LST) was first established by correlating it with ground-based measurements of Heat Index, as supplied by the PAGASA through their ground stations in NCR. The validation process involved utilizing two statistical techniques: Ordinary Least Squares (OLS) and General Linear Regression (GLR). These methods were employed to acquire the correlation coefficient (r-squared) to assess the consistency and reliability of LST measurements by comparing them with the observed Heat Index data collected from the ground.

Air quality could be determined by the concentration of air pollutant parameters within the ambient study area. The following air pollutants assessed in this study are CO, NO_2, and SO_2 (mol/m²). Predominant sources of these pollutants are combustion of carbonaceous fuels often found in vehicles for CO, and intense temperature combustion of fuels in domestic and industrial processes for NO_2 and SO_2. The density of air pollutants determines the Air Quality Index (AQI). The World Health Organization (WHO) has provided guidelines on tolerable values of these pollutants to lessen the risk they pose to the community.

2.3 GIS Regression Analysis

Generalized Linear Regression (GLR) and Ordinary Least Squares (OLS) were further used to assess the correlation between LST and air pollutants. LST was the dependent variable whereas CO, NO_2, and SO_2 are the explanatory variables. Subsequently, the correlation of the DEM on these variables was assessed, with DEM as the dependent variable and LST, CO, NO_2, and SO_2 as explanatory variables. The regression analysis was performed on different time frames: (1) from January 2019 to March 2023, (2) yearly, and (3) entire wet and dry seasons from January 2019 to March 2023.

The R-squared value, also known as the coefficient of determination, represents the percentage of variance in the dependent variable that can be explained by the independent variable (Moore, et.al., 2013). To interpret the strength of a relationship based on its R-squared value, these are the general guidelines in interpreting R-squared: the correlation below 0.3 is none to very weak; between 0.3 and 0.5 is weak; 0.5 to 0.7 is moderate; above 0.7 is deemed strong.

Aside from R-Squared, GLR and OLS shows other statistics that could determine the significance and relationship between the variables such as Pearson Correlation, and probability from Joint-F statistic, Joint-Wald Statistic, Koenker Statistic, and Jarque-

Bera statistic, and the intercept and coefficient of the linear equation formed from the regression (ArcGIS Pro, 2023). Joint-F and Joint-Wald Statistic indicate the overall model significance and may only be used when the Koenker statistic is not statistically significant. Koenker static determines the consistency relative to the spatial characteristics of the exploratory variables such that if deemed statistically significant, there may be non-stationarity or heteroscedasticity in the data. In addition, if Koenker is statistically significant, the Robust Probability should be used to determine the significance of the coefficient. Jarque-Bera statistics assessed the bias in the regression models, on which the residuals are not normally distributed. The probability of these statistics was named as *p(F)*, *p(K)*, *p(W)*, *p(RC)*, and *p(J)* for Joint F-statistic, Koenker statistic, Joint-Wald statistic, Coefficient Robust Probability, and Jarque-Bera statistic, respectively. If these statistics were determined to be statistically significant ($p<.01$), the regression may not be reliable.

3. RESULTS AND DISCUSSION

3.1 Correlation of LST with Heat Index

In Table 1, the OLS results show a strong correlation between mean LST and mean Heat Index, with a statistically significant R^2 of 0.803. Subsequently, this is consistent with GLR results ($R^2 = 0.81$). Both the regression models for the mean and maximum values display statistically insignificant Koenker and Jarque-Bera statistics ($p>.01$). Hence, these models are consistent with their spatial relationship and are free of bias. Consequently, the Joint-F statistic, Joint-Wald statistic, and Coefficient Robust Probability were all statistically significant, thus the independent variables improved the fit of the data in the regression model.

This reinforces the idea that changes in LST are closely associated with variations in Heat Index, suggesting a strong connection between the two (shown in Figure 2). Thus, LST from satellite data may be utilized in exploring relationships with air quality parameters.

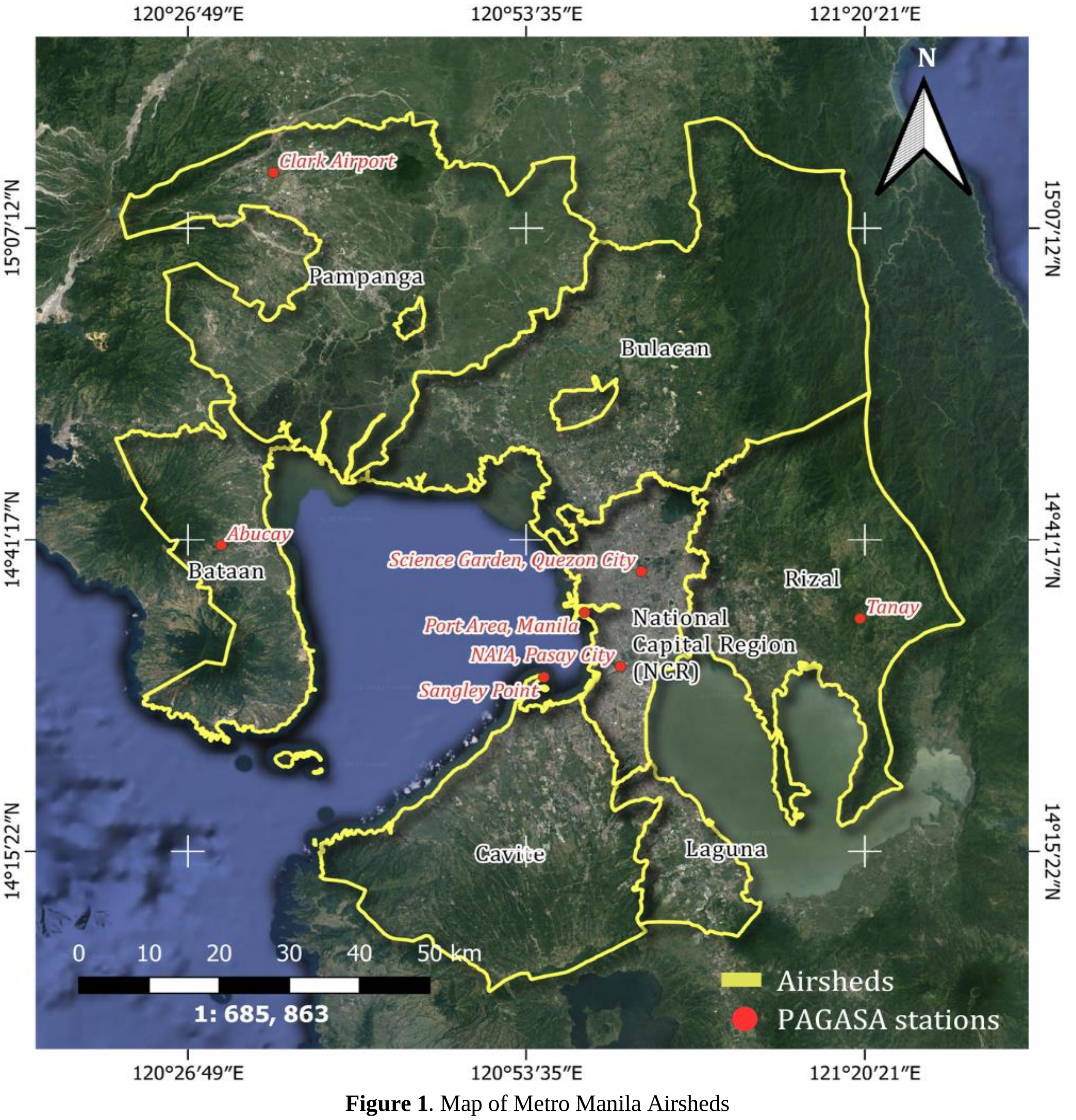

Figure 1. Map of Metro Manila Airsheds

Linear Regression	Model Summary							Parameter Estimates	
	R	*R²*	*p(F)*	*p(K)*	*p(W)*	*p(R)*	*p(J)*	Intercept	Coefficient
LST – HI Mean	0.896	0.803	.000	0.246	.000	.000	0.424	4.487	0.930
LST – HI Maximum	0.656	0.431	<0.001	0.629	.000	.000	0.200	11.726	0.757

Table 1. Linear Regression among mean and maximum of LST and Heat Index.

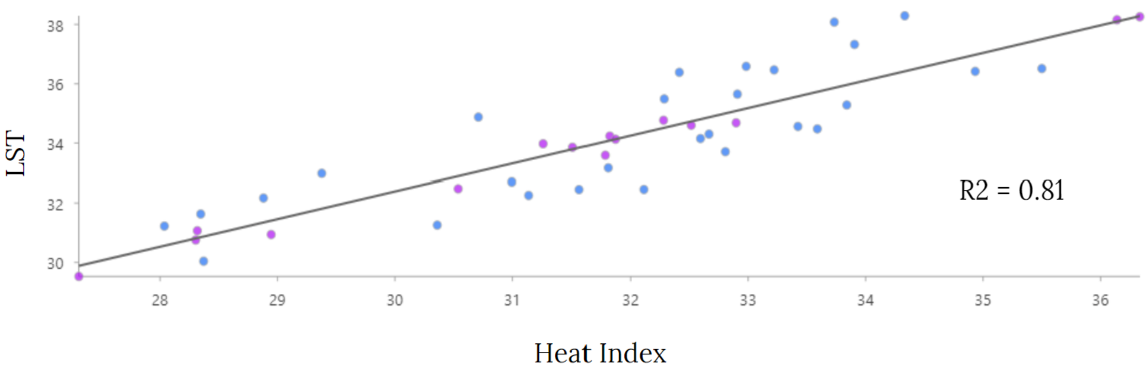

Figure 2. Correlation between Mean Land Surface Temperature and Mean Heat Index

3.2 Regression Model Analysis

Results from the models have shown a none to very weak correlation as shown in Table 2. There have been few cases of weak to low correlation mostly found between DEM and LST, LST and NO_2, and LST and CO.

The regression models between DEM to LST, CO, and NO_2 were found to have a negative correlation. Hence, the higher the elevation, the lower temperature, and the concentration of CO and NO_2 (shown in Figure 3). On these models, the highest correlation and coefficient of determination was observed between DEM and LST in 2022 data (R = -0.663, R^2 = 0.439). The lowest correlation and coefficient of determination was found between DEM and CO in 2021 data (R = -0.162, R^2 = 0.026).

In addition, DEM had almost no correlation with SO_2 concentrations (R = 0.000) with also having zero coefficient of determination. Compared to other explanatory variables, SO_2 is scattered across the study area (shown in Figure 3). Therefore, the elevation of an area does not provide meaningful information on the presence of SO2 pollutants in that area.

LST was known to be positively correlated with CO and NO_2 (i.e., the higher the land surface temperature, the higher the CO and NO_2 concentrations in area). Among these models, the highest correlation and coefficient of determination was seen between LST and NO_2 during the aggregated wet periods from January 2019 to March 2023 (R = 0.619, R^2 = 0.383). Subsequently, the lowest correlation and coefficient of determination on these models was between LST and CO during 2021 (R = 0.234, R^2 = 0.055). This positive weak correlation between LST and NO_2 was found consistent with the findings of the effect of LST on NO_2 concentration in Delhi, India using Google Earth Engine during summer and winter seasons from 2019 to 2021 (Rahaman et al., 2023). Furthermore, all highest LST and concentration of CO and NO_2 were seen clustering in the urban areas of Metro Manila (shown in Figure 3). These may indicate that the urban landscape of Metro Manila may have contributed to these high values.

For LST and SO_2, there was almost no correlation found with highest coefficient of determination in the 2019 and 2020 datasets (R^2 = 0.009) and lowest in 2021 datasets (R^2 = 0.001). This shows that SO_2 and LST do not provide any association, such that removing SO_2 from the model of LST would not result in any change. This is consistent with the relationship between DEM and SO_2, and DEM and LST.

Despite the relationship and the respective correlation coefficient of these regression models, issues on other statistics were seen that would lead to their weak reliability. For their Koenker statistics, only three regression models were found to be statistically insignificant — DEM and SO_2 in 2020 (p=.014), and DEM and SO_2 in aggregated dry seasons (p=.345), and LST and CO in covered 2023 (p=.012). This indicates that aside from these models, there is inconsistency with the spatial relationship of the variables such that there may be a presence of non-stationarity and heteroscedasticity. On these statistically significant Koenker statistics values, their respective Joint-Wald and Joint-F statistics values were used to determine the overall significance of the model. Joint-F and Joint-Wald statistics values mostly had a p=.000. This is less than the confidence level threshold of p<.01.

These findings were consistent with the results found in examining the relationship between the same concerned variables in Tehran, Iran (Fuladlu & Altan, 2021). Although the study had the same Joint-F statistical significance of zero on almost all their regression models, they were still reported reliable to infer the relationship between the variables.

In addition to these models' statistical significance of p=.000, the Coefficient of Robust Probability and the Jarque-Bera statistic also had the same significance. Although there were minimal cases of regression models having not exactly zero probability (i.e., p<.001), these were still considered statistically significant with the set confidence level threshold of p<.001. This posits that there is bias in these regression models, and they are deemed unreliable for any future references or policy making. This could be due to a misspecification, missing variables, or few samples in the regression models (*How OLS Regression works—ArcGIS Pro | Documentation*, n.d.). Furthermore, this exact zero probability may have introduced overfitting or underfitting in the regression models.

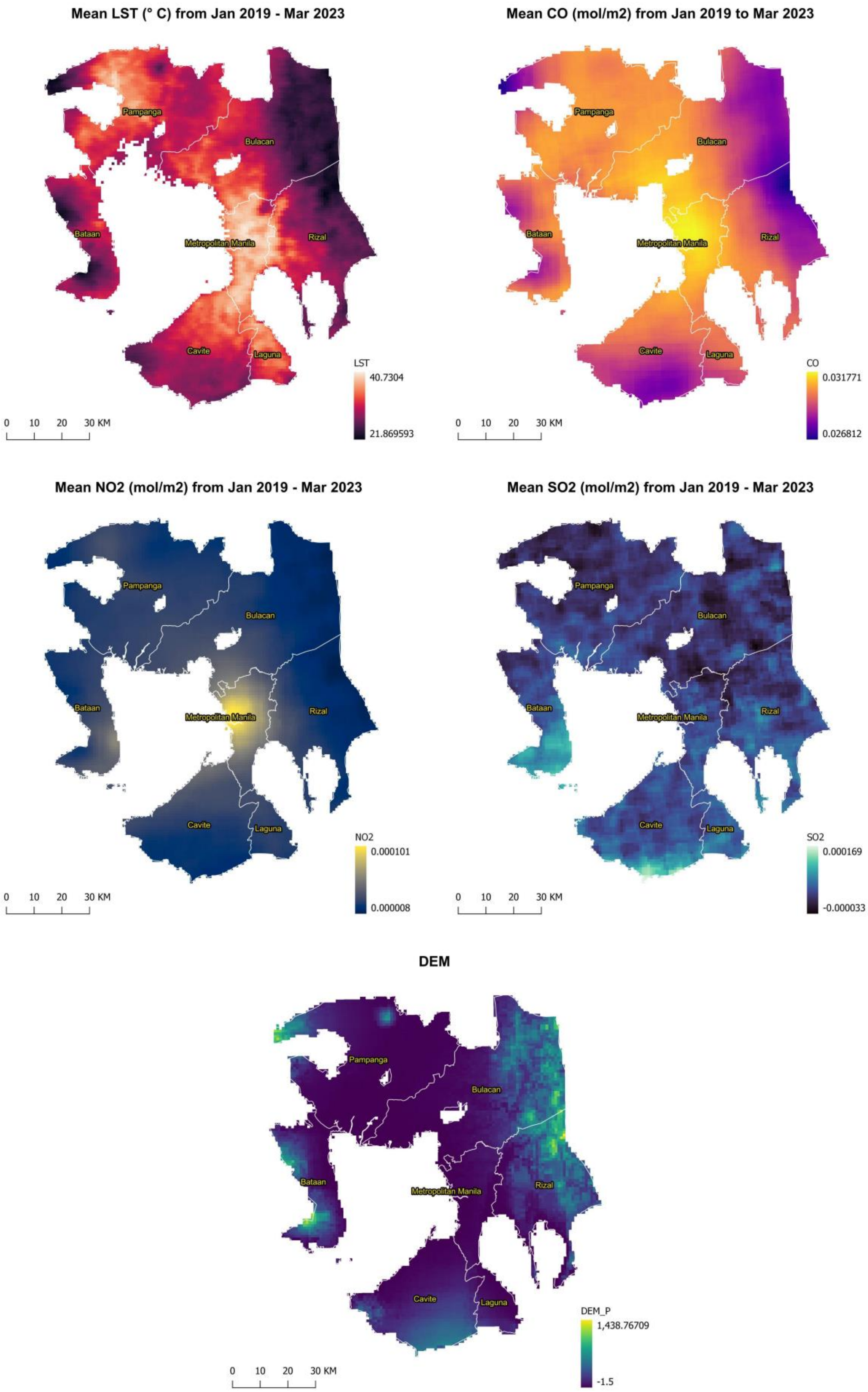

Figure 3. Mean data from January 2019 to March 2023

Period	Linear Regression	Model Summary							Parameter Estimates	
		R	***R^2***	***p(F)***	***p(K)***	***p(W)***	***p(RC)***	***p(J)***	**Intercept**	**Coefficient**
2019	DEM - LST	-0.610	0.372*	.000	.000	.000	.000	.000	1084.162	-29.496
	DEM - CO	-0.309	0.095	.000	.000	.000	.000	.000	498.656	-10, 868.265
	DEM - NO_2	-0.377	0.142	.000	.000	.000	.000	.000	302.947	-4, 663, 037.420
	DEM - SO_2	0.046	0.002	.000	.000	.000	.000	.000	172.677	113, 645.714
	LST - CO	0.352	0.075	.000	.000	.000	.000	.000	21.183	601.092
	LST - NO_2	0.584	0.341*	.000	.000	.000	.000	.000	26.371	151, 619.259
	LST - SO_2	-0.097	0.009	.000	<.001	.000	.000	.000	30.701	-5, 289.455
2020	DEM - LST	-0.602	0.362*	.000	.000	.000	.000	.000	1050.729	-28.512
	DEM - CO	-0.170	0.029	.000	.000	.000	.000	.000	422.296	-8, 226.206
	DEM - NO_2	-0.402	0.161	.000	.000	.000	.000	.000	321.944	-597, 496.314
	DEM - SO_2	0.010	0.000	<.001	.014**	<.001	<.001	.000	171.935	24, 220.170
	LST - CO	0.300	0.090	.000	.000	.000	.000	.000	21.114	323.304
	LST - NO_2	0.517	0.267	.000	.000	.000	.000	.000	26.885	164, 604.054
	LST - SO_2	-0.094	0.009	.000	.000	.000	.000	.000	31.040	-4, 764.134
2021	DEM - LST	-0.599	0.359*	.000	.000	.000	.000	.000	1000.211	-27.115
	DEM - CO	-0.162	0.026	.000	.000	.000	.000	.000	391.851	-7, 224.416
	DEM - NO_2	-0.416	0.173	.000	.000	.000	.000	.000	335.367	-6, 039, 472.130
	DEM - SO_2	-0.017	0.000	.000	.000	<.001	<.001	.000	169.191	-27, 575.146
	LST - CO	0.234	0.055	.000	.000	.000	.000	.000	23.900	218.986
	LST - NO_2	0.602	0.362*	.000	.000	.000	.000	.000	25.326	190, 962.855
	LST - SO_2	-0.034	0.001	.000	.000	.000	.000	.000	30.602	-1, 220.198
2022	DEM - LST	-0.663	0.439*	.000	.000	.000	.000	.000	1, 149.129	-31.332
	DEM - CO	-0.250	0.063	.000	.000	.000	.000	.000	623.881	-16, 044.447
	DEM - NO_2	-0.416	0.173	.000	.000	.000	.000	.000	321.071	-5, 228, 901.766
	DEM - SO_2	0.014	0.000	<.001	.000	<.001	<.001	.000	170.515	21, 567.153
	LST - CO	0.304	0.093	.000	.000	.000	.000	.000	19.160	404.545
	LST - NO_2	0.540	0.291	.000	.000	.000	.000	.000	26.516	146, 603.463
	LST - SO_2	-0.055	0.003	.000	.000	.000	.000	.000	30.679	-1, 846.230
2023 (Jan - Mar)	DEM - LST	-0.655	0.430*	.000	.000	.000	.000	.000	1, 095.157	-32.195
	DEM - CO	-0.346	0.120	.000	.000	.000	.000	.000	906.408	-22, 776.411
	DEM - NO_2	-0.327	0.107	.000	.000	.000	.000	.000	301.525	-4, 661, 288.973
	DEM - SO_2	0.062	0.004	.000	.000	.000	.000	.000	165.308	133, 467.147
	LST - CO	0.602	0.362*	.000	<.001	.000	.000	.000	2.486	806.526
	LST - NO_2	0.572	0.327	.000	.000	.000	.000	.000	23.743	168, 362.045
	LST - SO_2	0.045	0.002	.000	.010	.000	.000	.000	28.330	1, 990.061
Jan 2019 - Mar 2023	DEM - LST	-0.613	0.375*	.000	.000	.000	.000	.000	1, 051.319	-28.512
	DEM - CO	-0.204	0.042	.000	.000	.000	.000	.000	919.227	-24, 406.095
	DEM - NO_2	-0.395	0.156	.000	.000	.000	.000	.000	316.849	-53, 195, 953.907
	DEM - SO_2	0.012	0.000	.000	<.001	.000	.000	.000	170.578	21, 350.070
	LST - CO	0.274	0.075	.000	.012**	.000	.000	.000	21.154	311.787
	LST - NO_2	0.549	0.302*	.000	.000	.000	.000	.000	26.218	159, 921.095
	LST - SO_2	-0.065	0.004	.000	.000	.000	.000	.000	30.631	-2, 651.818
Wet Season (Jan 2019 - Mar 2023)	DEM - LST	-0.632	0.399*	.000	.000	.000	.000	.000	1, 140.488	-30.587
	DEM - CO	-0.198	0.039	.000	.000	.000	.000	.000	510.386	-12, 269.264
	DEM - NO_2	-0.376	0.141	.000	.000	.000	.000	.000	308.910	-5, 291, 407.162
	DEM - SO_2	0.012	0.000	.000	<.001	<.001	<.001	.000	170.833	20, 171.581
	LST - CO	0.358	0.128	.000	.000	.000	.000	.000	18.823	454.259
	LST - NO_2	0.619	0.383*	.000	.000	.000	.000	.000	26.596	165, 476.593
	LST - SO_2	-0.074	0.006	.000	.000	.000	.000	.000	31.478	-2, 641.939
Dry Season (Jan 2019 - Mar 2023)	DEM - LST	-0.612	0.375*	.000	.000	.000	.000	.000	1, 024.789	-28.304
	DEM - CO	-0.288	0.083	.000	.000	.000	.000	.000	776.781	-18, 812.180
	DEM - NO_2	-0.420	0.176	.000	.000	.000	.000	.000	328.331	-5, 433, 363.623
	DEM - SO_2	0.011	0.000	<.001	.345**	<.001	<.001	.000	170.324	23, 094.965
	LST - CO	0.495	0.245	.000	.000	.000	.000	.000	7.916	681.400
	LST - NO_2	0.477	0.227	.000	.000	.000	.000	.000	26.082	147, 383.860
	LST - SO_2	-0.054	0.003	.000	.000	.000	.000	.000	29.910	-2, 437.799

Table 2. Ordinary Least Squares (OLS) Regression results (*** =** *weak correlation*, ****** *$p > .01$)*

4. CONCLUSION

This study examines the correlation between land surface temperature, elevation, and air quality parameters (CO, NO_2, and SO_2) in the Metro Manila Airshed under different periods from January 2019 to March 2023. Using Ordinary Least Squares and Generalized Linear Regression, LST and heat index had a strong correlation, allowing the use of LST for further analysis.

Across different combinations between LST, elevation, and air quality parameters, a weak to low negative correlation was seen between DEM to LST, CO, and NO_2. In addition, weak to low positive correlation was seen between LST to CO and NO_2. Almost no correlation was found between DEM and SO_2, and LST and SO_2. However, these models might have the presence of overfitting or underfitting, non-stationarity or heteroscedasticity, bias, or misspecification as implied by their statistical significance across different statistical parameters.

The values of LST, CO, NO_2 were found to be clustered in the urbanized area of Metro Manila. Since this is a common observation among the variables, the urban landscape of an area may have contributed to the concentration of these pollutants. Therefore, it is strongly recommended that future research investigate factors influencing the significance of the correlation between LST and air quality parameters. Utilizing appropriate indices for both LST and various air pollutants will establish an extended systematic analytical framework. Notably, the Normalized Difference Built-Up Index (NDBI) and the Normalized Difference Vegetation Index (NDVI) are vital indices that play a pivotal role in comprehending the relationship and patterns between LST and air pollutants. NDBI helps identify areas where high urbanization can influence air pollutant concentrations, while NDVI emphasizes the cooling and air quality advantages of vegetation. Both indices contribute to an inclusive understanding of how land use, urbanization, and vegetation collectively influence the interrelationship between land surface temperature and air pollutants. Furthermore, investigating other pollutants such as particulate matter as an air quality parameter is recommended to shed more understanding on the relationship of land surface temperature and air quality.

Moreover, the use of other GIS-based regression analysis is suggested to determine the best-fit model for these variables. As seen by their statistics, additional pre-processing and training on the models would improve accuracy. Application of 70-30 regression partition for training and test sets through machine learning for modeling linear regression is recommended. Aside from other regression models, it is also recommended for these datasets to be assessed in other aspects such as autocorrelation on the residuals of the regression model, and spatiotemporal hotspot analysis to identify any potential effect they had on the fit of the model. Due to the large dataset, processing faced issues on hardware, hence a stronger machine is required to perform all processes and recommendations.

Nonetheless, it is advised for authorities to implement policies and actions to improve the status of air quality in Metro Manila airsheds. It must be ensured that despite the goals and actions towards urbanization, there should be utmost priority and balance for sustainability and environmental management, not only for the localized area but for the larger contribution it has on global climate change.

REFERENCES

ArcGIS Pro, 2023. How OLS regression works—ArcGIS Pro | Documentation. pro.arcgis.com/en/pro-app/ latest/tool-reference/spatial-statistics/how-ols-regression-works.htm arc

Copernicus Sentinel-5P (processed by ESA), 2020, TROPOMI Level 2 Sulphur Dioxide Total Column. Version 02. European Space Agency. doi.org/ 10.5270/S5P-74eidii

Copernicus Sentinel-5P (processed by ESA), 2021, TROPOMI Level 2 Carbon Monoxide total column products. Version 02. European Space Agency. doi.org/10.5270/S5P-bj3nry0

Copernicus Sentinel-5P (processed by ESA), 2021, TROPOMI Level 2 Nitrogen Dioxide total column products. Version 02. European Space Agency. doi.org/10.5270/S5P-9bnp8q8

Environmental Management Bureau, 2019. EMB holds airshed capacity building in emission inventory and air quality monitoring. air.emb.gov.ph/emb- holds-airshed-capacity-building-in-emission-inventory-and-air-quality-monitoring/

Fuladlu, K., & Altan, H., 2021. Examining land surface temperature and relations with the major air pollutants: A Remote sensing research in case of Tehran. *Urban Climate*, 39, 100958. doi.org/10.1016/j.uclim .2021.100958

Moore, D.S., Notz, W., Fligner, M.A., 2013., in: The Basic Practice of Statistics: David S. Moore, William I. Notz, Michael A. Fligner. W.H. Freeman and Co., New York, p. 138.

Official Gazette of the Republic of the Philippines, 2011. DENR to enhance air quality management program | GOVPH. www.officialgazette .gov.ph/2011/10/22/denr-to-enhance-air-quality-management-program/

Rahaman, S. N., Ahmed, S. M., Zeyad, M., & Zim, A. H., 2023. Effect of vegetation and land surface temperature on NO2 concentration: A Google Earth Engine-based remote sensing approach. *Urban Climate*; Elsevier BV. doi.org/10.1016/j.uclim.2022.101336

Wan, Z., MODIS Web, 2013. MODIS Land Surface Temperature and Emissivity (MOD11). modis.gsfc.nasa.gov/ data/dataprod/mod11.php

Wang, Y., Guo, Z., & Han, J., 2021. The relationship between urban heat island and air pollutants and them with influencing factors in the Yangtze River Delta, China. *Ecological Indicators*, 129, 107976. doi.org/10.1016/j.ecolind.2021.107976

World Health Organization, 2021. WHO global air quality guidelines: particulate matter (PM2.5 and PM10), ozone, nitrogen dioxide, sulfur dioxide and carbon monoxide www.who.int/publications/i/item/9789240034228