The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Volume XLVIII-5/W4-2025
Philippine Geomatics Symposium (PhilGEOS) 2025 "Enhancing Human Quality of Life through Geospatial Technologies",
24–25 November 2025, Quezon City, Philippines

Mapping the Unseen: The Impact of Urban Sprawl on NO2, CO, and O3 Pollutants in Cavite Using Sentinel-2 and Sentinel-5P from 2020 to 2023

Jeram Evered N. Colendres [1], Jane Camille N. de Vera [1], Mikaela P. Santos [1], Jommer M. Medina [1]

[1] Department of Geodetic Engineering, University of the Philippines, Diliman, Quezon City, Philippines –
(jncolendres, jndevera, mpsantos17, jmmedina)@up.edu.ph

Keywords: Land Cover Change, Air Pollution, Google Earth Engine (GEE), Supervised Classification, ANOVA Test

Abstract

This study utilizes remote sensing technologies and the Google Earth Engine (GEE) platform in assessing the impacts of urban sprawl on air quality in Cavite for 2020-2023. Sentinel-2 L2A images were used to generate 10 m resolution land cover classification map using supervised classification technique applied on training data for built-up, bare soil, vegetation water, and clouds for all the years. Accuracy assessment showed an overall accuracy of at least 99.90% across the years, with the κ ranging from 0.98 to 0.99. The rate of urbanization was found to increase at least 1% every year, from 25.04% in 2020 to 28.92% in 2023. For air quality, NO2, CO, and O3 derived from Sentinel-5P data for the same period were aggregated to quarterly values to address data gaps. Time-series analysis of the air quality parameters showed that NO2 peaked in the fourth quarter due to cooler weather, CO spiked in the second quarter from energy used during dry season, while O3 peaked in the second and third quarters due to higher temperature. Single-factor ANOVA test ($\alpha = 0.05$) showed that the NO2 and CO concentrations are significantly different between built-up and non-built-up areas, while O3 showed no significant difference. Regression analysis also showed that NO2 and CO concentrations increase with the expansion of built-up areas, whereas O3 was not directly affected.

1. Introduction

1.1 Background of the Study

Urban sprawl happens when a city expands to the lower density areas due to the increase in population causing the dispersion of residents (National Geographic Society, 2023). In Cavite, an average annual urbanization rate of 6.47% was recorded from the years 2015 to 2020 (Cavite | Ecological Profile, 2020). The growth of residential neighbourhoods and the entry of businesses are responsible for the significant increase in the urban population in urban barangays (Cavite | the Official Website of Cavite, 2024). While urbanization continues to drive economic growth, it also brings challenges like environmental degradation and rising pollution levels, particularly affecting air quality.

Urbanization in the Philippines has led to increased vehicular traffic as well as industrial and commercial activities. In the Philippines, especially in parts of Cavite, air pollution has worsened due to these factors, including use of fossil fuels, emissions from households and from private and public transportation. Vehicles emit high levels of pollutants such as carbon monoxide (CO) and nitrogen oxides (NO) (Brooke, 2024). In Cavite, emissions from transportation and vehicles are the main sources of air pollutants in the city (IQAir, 2024). This situation underscores the urgent need for close monitoring of air quality, especially in areas where urban sprawl is most evident.

However, the challenge lies in the limited availability and spatial coverage of ground-based air quality sensors, which are often too sparse to capture local variations in pollution levels. To address this, remote sensing-derived air quality data offers a valuable alternative, providing broader spatial coverage and enabling more detailed analysis of the relationship between land cover and air quality parameters which is an issue that is particularly relevant in peri-urban areas experiencing urban sprawl, such as Cavite. With this, the researchers intend to determine the growth in urban area through level 1 supervised land cover classification. In addition, the researchers aimed to study the impact of urban sprawl on air quality specifically, Nitrogen dioxide (NO2), Carbon monoxide (CO), and Ozone (O3) in Cavite for the years 2020 to 2023.

1.2 Research Objectives

The study aims to utilize remote sensing data and techniques to assess the impacts of urban sprawl on air quality in Cavite. Specifically, it also has these following objectives:

1. Determine the rate of urbanization in Cavite using Sentinel-2-derived land cover maps for 2020-2023;
2. Compare the atmospheric concentrations of NO2, CO, and O3 in built-up and non-built-up areas in Cavite from 2020 to 2023 using Sentinel-5P; and
3. Examine the relationship between land cover types and air quality parameters using regression analysis.

1.3 Significance of the Study

The findings of this study align the Sustainable Development Goals (SDGs). In particular, the study supports SDG 11: Sustainable Cities and Communities especially Target 11.3, which emphasizes enhancing inclusive and sustainable urbanization and Target 11.6, which calls for reducing the adverse environmental impact of cities. Furthermore, the study contributes to SDG 3: Good Health and Well-Being, specifically Target 3.9, which aims to significantly decrease health issues resulting from pollution in air. Understanding the relationship between land cover and air quality offers a scientific foundation for strategies to safeguard public health as air pollution is a significant risk factor for respiratory and cardiovascular conditions.

2. Methodology

2.1 Study Area

The primary area of the study was the province of Cavite which consists of 7 cities and 16 municipalities as shown in Figure 1. As per the 2020 Census, Cavite has a total population of 4,344,829, rendering it the most densely populated province, covering an area of around 1,526 square kilometers (km^2), which results in a significant average density of roughly 2,847 individuals per km^2. Within its boundaries, the northern parts of Cavite which are closer to Metro Manila, which includes Rosario, GMA, Bacoor, Imus, Dasmariñas, and Cavite City, are classified as high-density urban areas, with Rosario boasting 19,543 per km^2. Surrounding these are medium-density area, which include Mendez, Silang, Naic, and Tagaytay City, all of which have population densities surpassing 1,000 individuals per km^2. On the other hand, the upland and mountainous municipalities, such as Maragondon and Magallanes, remain characterized by low population density and are primarily rural, with densities significantly below 1,000 individuals per km^2 (Cavite Ecological Profile, 2021).

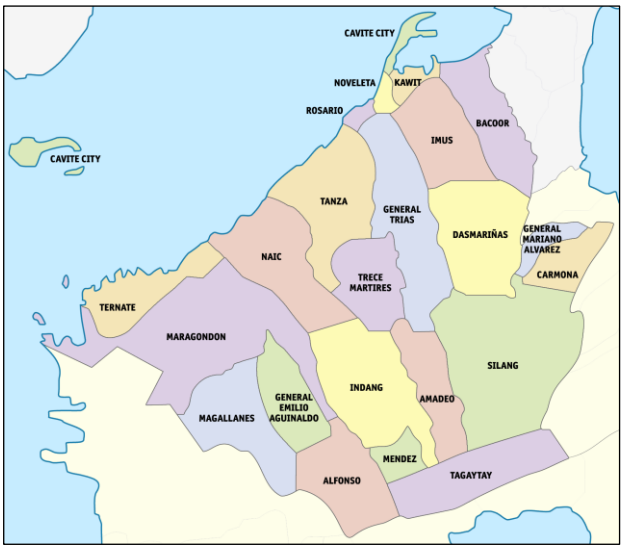

Figure 1. Study area (Proud Caviteño, 2019).

2.2 Datasets Used

Table 1 shows the data for this study which was gathered from Google Earth Engine (GEE). For the land cover classification, the Sentinel-2A satellite images for years 2020-2023 were extracted with a spatial resolution of 10 meters. Moreover, quarterly data for years 2020-2023 of NO2, CO, and O3 in Cavite were extracted using the Sentinel-5P satellite images.

Dataset	Spatial Resolution	Period of Coverage	Sources
Sentinel-2A	10 meters	2020-2023	GEE
Sentinel-5P	10 meters	2020-2023	GEE

Table 1. Summary of datasets used.

2.3 Data Processing

Figure 2 displays the general workflow of the methodology of the study. The procedure began with collection of data needed. Then, land cover classification together with accuracy assessment were done using the Sentinel-2A images and extraction of air quality pollutants was done using the Sentinel-5P images. To determine the relationship between the two, ANOVA test and Multiple Linear Regression test were conducted.

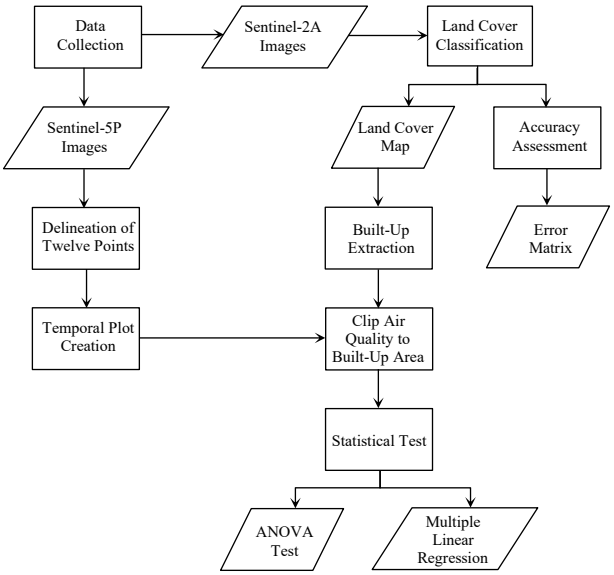

Figure 2. Flowchart of the methodology.

2.3.1 Land Cover Classification: Sentinel-2 satellite images was used for land cover classification as it provides more accurate results compared to Landsat-8 since it can identify smaller elements with greater accuracy due to its enhanced spatial resolution and improved spectral resolution of the red edge bands (Nwagoum, et al., 2023). Moreover, bottom-of-atmosphere (L2A) imagery derived from top-of-atmosphere (L1C) was used. The choice of L2A is supported by a study by Ginting et al. (2024), which reveals that it surpasses L1C. This is because L2A effectively addresses atmospheric, terrain, and cirrus effects on L1C data, in addition to having undergone radiometric correction.

2.3.1.1 Land Cover Classification Process: In this study, the satellite imagery used covered an annual cycle, with the datasets selected aligning with mid-year acquisitions, which consistently took place in the final two weeks of June. Four images were processed for land cover mapping. The time frame from 2020 to 2023 was deliberately selected, as 2020 marks the latest dataset available from Cavite EP, while 2023 was set as the final year to ensure complete coverage, given that the study was conducted in the last quarter of 2024.

Level 1 land cover classification map was then created by creating training data for bare soil, built-up area, cloud, vegetation, and water. This was adopted following the approach of Kandel and Pokhrel (2024), as it provides a broad yet effective framework for analyzing urban sprawl. Training data were generated using Google Earth due to the infeasibility of fieldwork, with sufficient samples spread evenly for each class. These datasets were updated annually to reflect changes in land cover caused by factors such as cloud cover and urbanization.

Subsequently, supervised classification technique was applied on satellite images, employing the Maximum Likelihood Classification (MLC). MLC is superior and surmounts other algorithms used in land cover classification such as minimum distance and support vector machine (Dixit & Jog, 2016; Abassi et al., 2021). Both spatially and temporally, the approach used offers an effective tool for monitoring land cover change and for further analysis and interpretations (Olfato-Parojinog et al., 2023). Moreover, to enhance the result of the classification, the

study utilized Majority Analysis to remove single pixels that are surrounded by a large cluster of pixels of the same class.

2.3.1.2 Accuracy Assessment: Validation data for each year were generated using Google Earth imagery and constructed through a stratified sampling method to ensure an even distribution across the map and adequate representation of all land cover classes. Different sets were created annually to account for changes in land cover. Accuracy assessment was conducted using an error matrix to compute the User's, Producer's and Overall Accuracy. A minimum accuracy of 85% was set, following Kandel and Pokhrel (2024), to ensure the reliability of the classification.

2.3.1.3 Built-up Extraction: Extraction of built-up areas from the generated land cover classification was employed. The area of the built-up area extracted was computed each year to determine the rate of urbanization in Cavite. Limitations include possible errors in the land cover classification such as the area of the built-up area due to the presence of clouds.

2.3.2 Air Quality Parameters: Sentinel-5P was utilized in this study due to its reliable spatial resolution and its established application in measuring atmospheric air pollution parameters (Kaplan et al., 2019). A total of 48 satellite images were processed, with 16 images allocated for each 3 pollutants. The images were collected quarterly across the study period to minimize overgeneralized of values when averaged, thereby ensuring a more representative temporal analysis of pollutant concentrations.

The pollutants selected for this study were nitrogen dioxide (NO2), carbon monoxide (CO), and ground-level ozone (O3). The selection of NO2, CO, and O3 as pollutants of interest was based primarily on the availability of satellite imagery in GEE and their relevance to public health (Garajeh et al., 2023). NO2 was included due to its adverse effects on the respiratory system, while CO was considered because of its association with traffic congestion and its impact on oxygen transport within the body. O3 was also selected as it poses significant health risks when concentrations exceed established air quality standards.

Then, time-series profile of the air quality was conducted. Twelve-point features were used in temporal plot, consisting of three sites in both Bacoor and Cavite City to represent built-up areas, and three sites in both Ternate and Amadeo to represent non-built-up areas. Random selection was employed in the selection of three sites to reduce potential bias and to ensure that the chosen locations provide a representative basis for further analysis of air quality variations.

2.4 Analysis of Variance (ANOVA) Test

ANOVA test is a statistical method designed to determine whether there are statistically significant differences between the means of two or more groups. It works by comparing the variance between groups with the variance within groups. A significant result, indicated by an F-statistics and a p-value less than 0.05, leads to rejecting the null hypothesis of no difference between group means (Kenton, 2025).

ANOVA test was employed to determine if NO2, CO, and O3 concentrations differ significantly between built-up and non-built-up areas. Results were interpreted using the p-value and F-statistic, where rejecting the null hypothesis indicated significant differences; otherwise, failing to reject it suggested that pollutant levels were not strongly influenced by urbanization.

2.5 Regression Analysis

Multiple linear regression is a statistical tool that is used to model the relationship between a dependent variable and two or more independent variables. The equation used to understand the dependent variable using one or more independent variables has the form (Joshi, 2023):

$$Y = b1X1 + b2X2 + ... + bnXn, \quad (1)$$

where Y = dependent variable
$X1, X2, ... Xn$ = independent variables
$b1, b2, ..., bn$ = coefficients of independent variable

Regression equations will provide valid and accurate baseline projections providing all parameters are in similar study region, scale, and period. Applying them beyond this field runs the risk of erroneous inference and loss of accuracy (Luijken et al., 2019; Khurshid et al., 2024). Multiple linear regression was employed to further assess the influence of land cover types, built-up areas, water, and vegetation, on NO2, CO, and O3 concentrations. These land cover classes were chosen due to their relevance to pollutant dynamics, with built-up areas acting as emission sources, water influencing pollutant dispersion, and vegetation affecting pollutant concentrations. Study sites in Bacoor, Trece Martires, and Ternate represented high, medium, and low urban density, respectively, with land cover quantified as area percentages.

In the regression models, the target results are the coefficients of each land cover which indicates the direction and magnitude of influence of the independent variables and the R^2 values which measures the model fit. Moreover, p-values of independent variables whose values are less than 0.05 are considered statistically significant, indicating great influence on the pollutant concentration. Lastly, Akaike information criterion (AIC) is a metric for evaluating model fit that weighs the number of variables in the model against goodness of fit (Portet, 2020). The model is more accurate when the AIC value is smaller. The equation used for the AIC metric is given by:

$$AIC = 2k - 2\ln(L), \quad (2)$$

where AIC = Akaike information criterion value
k = number of independent parameters + 1
L = maximized value of likelihood functions (observations/residual)

3. Results and Discussion

3.1 Land Cover Classification

An assessment of land cover changes using supervised classification techniques from 2020 to 2023 revealed that Cavite exhibited a variety of features, including bare soil, built-up, vegetation, and water. Figures 3-6 show the land cover classification map in Cavite from 2020 to 2023, respectively.

Accuracy assessment results of 2020 to 2023 land cover maps showed a consistently high accuracy. The summary of results is shown in Table 2 showing the classes bare soil, built-up area, cloud, vegetation, and water, respectively. The results show that the overall accuracy remains above 99.90% across the four years, with the κ ranging from 0.98 to 0.99, demonstrating a strong alignment between the classified and validation data.

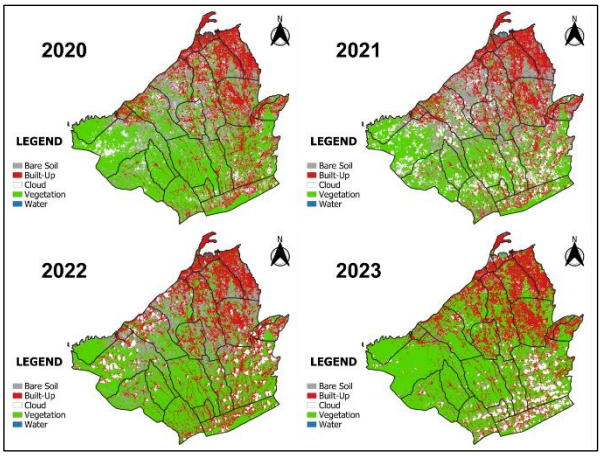

Figure 3. Land cover classification in Cavite for 2020-2023.

Notably, the Producer's Accuracy (PA) for built-up areas achieved a consistent 99% score for every year while the bare soil had 85.01% in 2023, but still within the set threshold of accuracy which was 85%. For built-up areas, the lowest and highest PA observed was 99.03% (2020) and 99.86% (2023), respectively. With respect to User's Accuracy (UA), most of the classes had values above 90% however, the Bare Soil had 88.46% in 2023. In addition, the lowest and highest UA observed for built-up areas was 99.03% and 99.86%, respectively.

Year	Producer's Accuracy (%)				κ
	Bare Soil	Built-Up	Veg	Water	
2020	98.35	99.03	99.95	99.94	0.98
2021	97.08	99.53	99.34	99.98	0.99
2022	98.15	99.71	99.76	99.93	0.98
2023	85.01	99.86	99.79	99.95	0.99

Year	User's Accuracy (%)				Overall Accuracy (%)
	Bare Soil	Built-Up	Veg	Water	
2020	100	92.48	98.83	100	99.94
2021	93.32	96.43	100	100	99.96
2022	98.19	94.7	99.95	100	99.92
2023	88.46	95.34	100	100	99.92

Table 2. Accuracy assessment results of the land cover maps.

From the generated area of built-up based on land cover classification, rate of urbanization was computed, as shown in Table 3 Cavite experienced an increase in urbanization of at least 1% or 15 sq. km every year from 2020 to 2023.

Year	Area Percentage	Rate of Urbanization
2020	25.04	(Base year)
2021	26.16	+1.12%
2022	27.91	+1.75%
2023	28.92	+1.01%

Table 3. Rate of urbanization in Cavite from 2020 to 2023.

The annual rate of urbanization reveals annual variations, signifying changes in the development pace. In 2021, the rate stood at 1.12%, marking a slight increase from the prior year. Notably, 2022 experienced the highest rate of 1.75%, indicating a rapid acceleration in urban growth, which is likely attributable to heightened construction activities and efforts aimed at economic recovery after the pandemic. Moreover, the urbanization rate reached 1.01% in 2023, suggesting a slowdown in urban growth, which may be influenced by restrictions on land availability or economic factors.

3.2 Quarterly Air Quality Concentration

The quarterly concentration of NO2, CO, and O3 in Cavite for the year 2023, as the other years yield similar maps, are shown in Figures 4, 5, and 6 respectively. The months of the year are represented by the initials JFM, AMJ, JAS, and OND.

In general, the concentrations were consistent for the same months on each year, further discussed on Section 3.3. In Figure 4, NO2 concentrations in the province of Cavite had higher concentrations on its upper part, near the boundary of Metro Manila, and lower on the lower regions, and built-up areas are denser in the upper regions compared to the lower region. CO concentrations also had similar results with NO2 as shown in Figure 5. Only O3 had a different result, with its concentrations being inconsistent all throughout, shown in Figure 6.

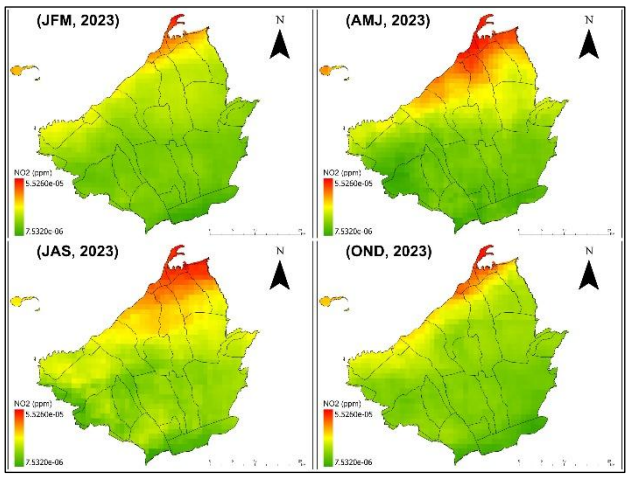

Figure 4. NO2 Concentration in Cavite in 2023.

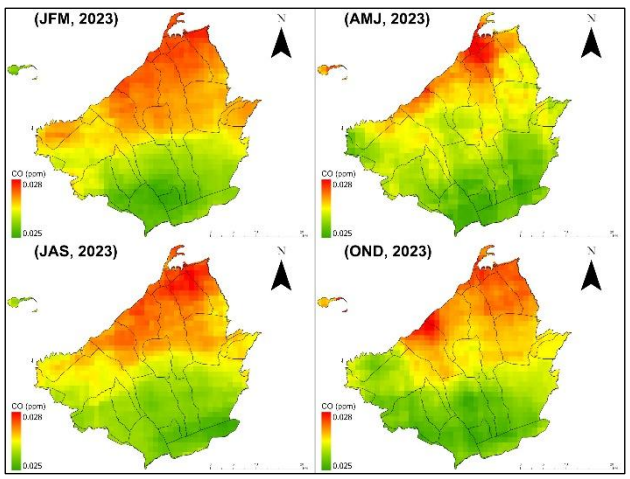

Figure 5. CO Concentration in Cavite in 2023.

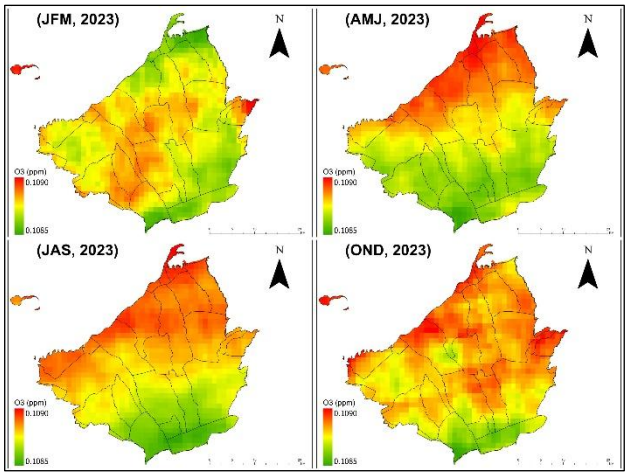

Figure 6. Ozone Concentration in Cavite in 2023.

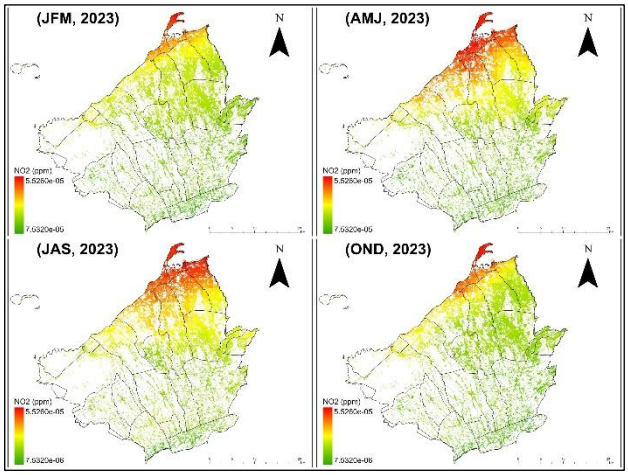

Figure 7. NO2 Concentration on Built-Up Areas in 2023.

Figure 7 shows that built-up areas classified have the highest concentration of the NO2 pollutants, indicated by the presence of the red color. It revealed that Bacoor, Cavite City, and other populated cities such as Dasmariñas as the area with the highest NO2 concentrations. On the other hand, Ternate and Amadeo have high constant concentrations of colors greens and yellows which indicates low concentrations of NO2. It can also be noted that the higher concentrations of NO2 can be seen on the upper portion of Cavite, near the border of National Capital Region (NCR) which can also be the reason of higher concentrations of NO2 clustering in the area. Therefore, it is reasonable to presume that there is a relatively large NO2 concentration in populated areas than those that are not as populated.

It is also evident that the amount of CO is relatively higher in built-up areas than those in forested areas or non-urban areas, as shown in Figure 8. The concentration values in Bacoor and Cavite City are consistently red all throughout, compared to those of Ternate and Amadeo which yield to yellow and green.

Moreover, Figure 9 shows that O3 concentrations are more spatially scattered compared to NO2 and CO, with values varying widely across locations. Areas like Bacoor and Cavite City display mixed concentration levels (red, green, and yellow), while parts of Ternante and Amadeo show highly concentrations (red), suggesting that O3 levels are not directly tied to population or built-up environment.

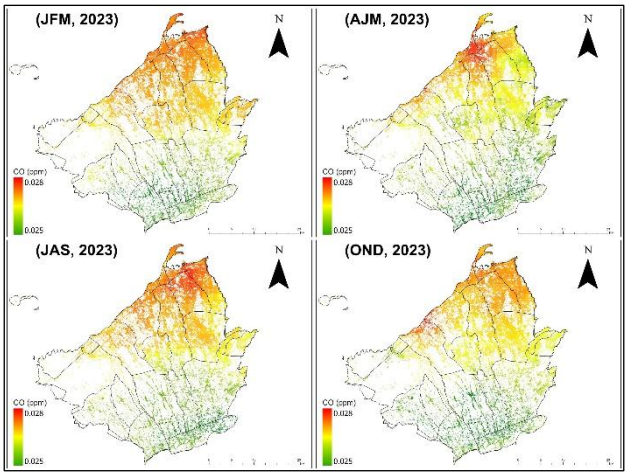

Figure 8. CO Concentration on Built-Up Areas in 2023.

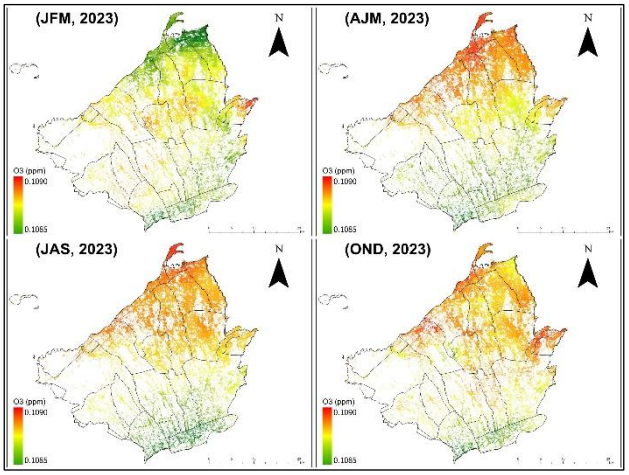

Figure 9. Ozone Concentration on Built-Up Areas in 2023.

3.3 Air Quality in Urban and Non-urban Areas

All the plots for each air pollutant were compiled into one temporal plot for comparison between urban and non-urban areas, represented by three locations for the two-each selected areas in Cavite: Bacoor, Cavite City & Ternate, Amadeo.

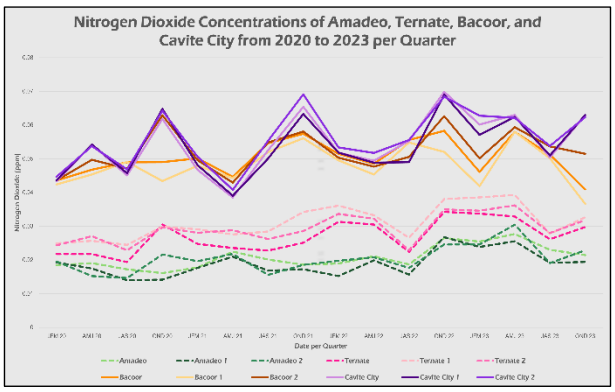

Figure 10. NO2 concentrations of twelve points in Cavite from 2020 to 2023.

In Figure 10, the peak of NO2 concentration happens during the 4th quarter which can be a result of the holiday season's increased demand for transportation, which raises fuel consumption. During the holidays, Grab regularly notices a sharp rise in demand for its mobility services. In December 2023,

GrabCar demand is predicted to rise by at least 20%, with a 45% increase in the second and third weeks (Roces, 2023).

Additionally, fireworks during holidays like Christmas and New Year's Eve release large quantities of dangerous chemicals like nitrogen oxides which exacerbates air pollution (Manila Observatory, 2022). Finally, cold air tends to trap contaminants closer to the ground because of its higher density, which lowers the quality of the air (Prana Air, 2024).

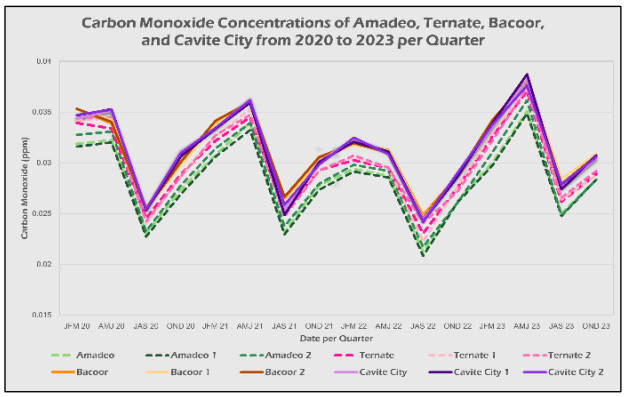

Figure 11. CO concentrations of twelve points in Cavite from 2020 to 2023.

Moreover, for CO pollutant, it is apparent from Figure 11 that the peak is during the second quarter which can be due to the increase in summertime electricity demand which leads to increased use of fossil fuels (IEA, 2023). Summertime travel may also rise because more people take trips and vacations, which raises car emissions. Additionally, less fuel efficiency results in higher fuel consumption because of the higher temperatures (Yateswrecker, 2024). Moreover, due to its lower density, hot air has less oxygen available for combustion than cold air. Fuel may burn less effectively as a result.

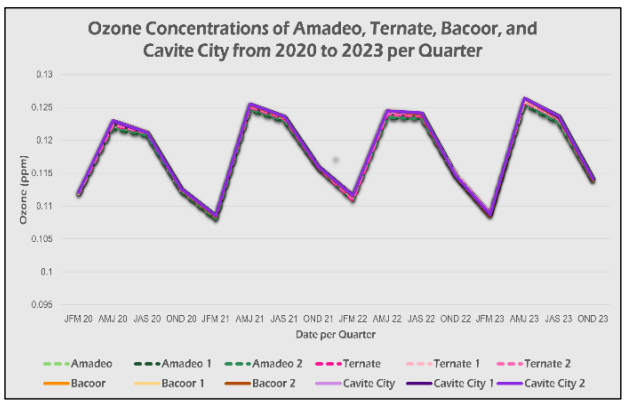

Figure 12. Ozone concentrations of twelve points in Cavite from 2020 to 2023.

Lastly, as shown from Figure 12, O3 concentrations peak at the second and third quarter which may be linked to summer temperatures rise as O3 respond to heat and sunlight (National Jewish Health, n.d.). The temperature has a great impact on ozone emergence, therefore the dry, hot months of the second quarter and the early, warm, humid months of the third quarter are optimal for its production. And it seems that ozone concentration does not depend on how populated an area is, but rather mostly on temperature fluctuation instead.

3.4 ANOVA Test Interpretation

Single-Factor ANOVA was used to determine whether the air quality concentrations, of NO2, CO, and O3, between urban and non-urban areas in Cavite are significantly different from one another, with a confidence level of 95% ($\alpha = 0.05$).

For NO2, results showed a highly significant difference between built-up and non-built-up areas, with a p-value of 5.17e-26 and an F-value (315.72) much greater than F-critical value (3.99). These findings indicate that there is a significant difference between NO2 concentrations in built-up and non-built-up areas. These further supports results previously discussed about the influence of built-up areas on the concertation of NO2 through urban activities especially during holidays.

Similarly, the analysis for CO showed a significant difference between built-up and non-built-up areas. The p-value (3.94e-2) was below the 0.05 threshold, and the F-value (4.43) exceeded the F-critical value (3.99), leading to the rejection of the null hypothesis. This result support the interpretation that CO levels are closely tied to urban-related activities, particularly in summer, when fossil fuels use and travels demands are higher, resulting in increased emissions.

In contrast, the results for O3 showed no significant difference between built-up and non-built-up areas. The p-value (0.79) was greater than 0.05, and the F-value (0.07) was below the F-critical value (3.99), indicating that the null hypothesis could not be rejected. This outcome suggests that O3 concentrations are not strongly dependent on urban sprawl but are instead can be influenced by factors such as temperature and sunlight, consistent with the patterns observed in the temporal plots.

3.5 Regression Analysis Interpretation

A total of 96 observations were analyzed, representing 73 barangays in Bacoor, 13 barangays in Trece Martires, and 10 barangays in Ternate, Cavite. Using regression analysis, the air quality concentrations served as the dependent variable, while land cover types were treated as the independent variable, with separate models developed for each air pollutant. Table 4 shows the regression model of each air pollutant has a form:

Y	Regression Model
NO2	5.85e-05 – 2.20e-07X1 – 3.88e-07X2 + 7.80e-07X3
CO	3.58e-02 – 6.43e-06X1 – 2.44e-05X2 + 1.37e-04X3
O3	1.14e-01 – 1.71e-06X1 – 1.28e-06X2 + 5.43e-06X3

Table 4. Summary of regression models.

where *Y* dependent variables are nitrogen dioxide, carbon monoxide, and ozone and *X1, X2,* and *X3* are the percentage of the total area of water, vegetation, and built-up land covers, respectively.

Regression Statistics		
Y	Adjusted R^2	AIC
NO2	25.56e-02	-2292.02
CO	28.34e-02	-1418.93
O3	6.26e-02	-1853.54

Table 4. Regression statistics summary output

From Table 4, the adjusted R^2 of 25.56% indicates that land cover explains a moderate portion of NO2 variation, showing that water, vegetation, built-up areas significantly influence its other environmental and anthropogenic factors. NO2 variability is complex since it is both a primary pollutant from thermal power plants and automobiles and a secondary pollutant formed through ozone, volatile organic compounds (VOCs), and other particulate matter (Wells, 2025), with seasonal fluctuations linked weather and traffic, especially during holidays. While land cover data cannot fully capture this temporal variability, the relationship remain meaningful: water cover is negatively and significantly associated with NO2 due to the absence of direct emissions, vegetation is also negatively and significantly related to NO2 as it offers reduction of NO2 levels through natural absorption, and built-up areas positively and significantly contribute to higher NO2 levels through industrial activities, combustion, and traffic (Borck & Schrauth, 2021).

Correspondingly, 28.34% of the variation in CO levels can be explained that carbon monoxide is strongly propelled by weather variability and other time-dependent processes, as also mentioned in the CO's temporal plot, due to the rise in the demand for electricity during the summer season (IEA, 2023). Among these, water cover showed a negative coefficient but statistically insignificant effect, indicating minimal influence on CO levels. Moreover, vegetation cover has a negative and significant coefficient, meaning that 1% increase in vegetation cover reduces CO by 2.44e-05 units, suggesting its role in mitigating air pollution. In contrast, built-up areas show a positive coefficient and significant relationship, indicating that more urbanized area directly increases CO concentrations, reinforcing the strong link between urban activities and elevated air pollution.

Relative to NO2 and CO, the adjusted R^2 for O3 concentration is notably low, reflecting its distinct behavior as a secondary pollutant formed through photochemical reactions that involves sunlight, VOCs, and nitrogen oxides (NO) (Yadav, 2023). This indicates that O3 levels are more strongly influenced by atmospheric and natural environmental processes such as solar radiation and other weather conditions, which land cover cannot adequately represent. Among the independent variables, water cover showed a negative but statistically insignificant coefficient, suggesting a slight moderating effect due to sea breezes that enhance dispersion of O3. Vegetation land cover also showed a negative but insignificant connection with ozone, highlighting the stronger role of broader atmospheric dynamics such as wind transport and intensity of solar radiation than vegetation alone. Lastly, built-up areas revealed a larger positive but insignificant coefficient. Overall, the results suggest that O3 concentrations in Cavite are not strongly determined by land cover classes and may be shaped by broader meteorological and climate variability.

4. Conclusion

Built-up area within the province of Cavite had increased by 3.88% from 2020 to 2023, indicating steady expansion of urban areas during the study period. This growth has induced anthropogenic activities causing emissions of certain air quality indicators into the atmosphere. Air quality maps generated from Sentinel-5P, showed that pollutant concentrations, particularly NO2 and CO, were higher in built-up areas compared to non-built-up areas. Nevertheless, there remain some built-up areas with relatively better air quality, reflecting that urbanization is not the sole factor influencing air pollution, as transportation, energy use, and fuel combustion also play major roles.

The statistical analyses further support these finding. Results from ANOVA test revealed that NO2 and CO levels in built-up areas were significantly higher compared to non-built-up areas, while O3 showed no significant difference. Regression analysis also indicated that NO2 and CO concentrations increase with the expansion of built-up land cover, whereas O3 was not directly affected by urbanization. These results further reinforce that carbon monoxide and nitrogen dioxide are closely tied to urban-related fuel consumption and emissions from vehicles, which is quite different from ozone that depends more on chemical reaction with volatile organic compounds, nitrogen oxides, temperature, and sunlight. Moreover, all independent variables are within the same context and same domain of data used such as similar environmental background and land-cover percentages. The time resolution and area of the input data correlate with that of the one used in the model calibration. With these, the regression equations are deemed valid and accurate, as also outlined by Luijken et al. (2019) and Khurshid et al. (2024) on predictor measurement and applicability domain.

Moving forward, further recommendations can be done to enhance the depth of the analysis in the study. It is recommended to conduct a multitemporal hot spot analysis for a deeper understanding of air quality trends in a spacetime context. It is also recommended to investigate particulate matter concentrations, PM2.5 and PM10, rather than ozone as these are the most monitored air quality parameters and considered critical due to their more direct and sever health impacts. Another recommendation is incorporate data from the ground-based air quality monitoring stations, nearest in Metro Manila, to improve the accuracy and reliability of the air quality data collected from Sentinel-5P as Cavite does not have any ground stations for air quality monitoring. Lastly, as the study currently only focuses on the correlation between urban sprawl and air pollutants, nitrogen dioxide, carbon monoxide, and ozone, in Cavite, it is recommended to include other factors for the pollution analysis such as climatological variables (e.g., wind current for dispersion of particles and exposure data (e.g., population) for the context of health impacts.

5. References

Abassi, H., Chungtai, A., Karas, I., 2021: A review on change detection method and accuracy assessment for land use land cover. R*emote Sensing Applications: Society and Environment*., 22, 2352-9385. doi.org/10.1016/j.rsase.2021.100482.

Borck, R., Schrauth, P., 2021: Population density and urban air quality. *Regional Science and Urban Economics*, 86, 103596, 0166-0462. doi.org/10.1016/j.regsciurbeco.2020.103596.

Brooke, E., 2024. Air pollution in the Philippines: Breathing towards a cleaner future. Breathe Safe Air. breathesafeair.com/air-pollution-in-the-philippines (23 November 2024).

Cavite | The official website of Cavite., 2024. cavite.gov.ph (24 August 2025).

Cavite Ecological Profile 2021 | Cavite., (2022). cavite.gov.ph/home/cavite-ecological-profile-2021/ (17 October 2024).

Dixit, M., Jog, S., 2016: Supervised classification of satellite images. *2016 Conference on Advances in Signal Processing (CASP)*, 93-98. doi.org/10.1109/CASP.2016.7746144.

Ginting, D., Setiawan, K., Anggraini, N., Suardana, A., Nandika, Ulfa, A., Aziz, K., & Dewanti, R., 2024: Comparison between top and bottom of atmosphere Sentinel-2 image for mangrove mapping in Balikpapan Bay, East Kalimantan. *BIO Web of Conferences,* 89, 07003. doi.org/10.1051/bioconf/20248907003.

IEA, 2023. Keeping cool in a hotter world is using more energy, making efficiency more important than ever – Analysis. iea.org/commentaries/keeping-cool-in-a-hotter-world-is-using-more-energy-making-efficiency-more-important-than-ever (21 November 2024).

IQAir, 2024. Philippines Air Quality Index (AQI) and Air Pollution information. iqair.com/philippines?srsltid=AfmBOooBvZSsTbdVpCTQ6T9dTlBMpk_YhPSmn5XytQsmXSqPbZFRnMBm (18 October 2024).

Joshi, B., 2023. Understanding simple linear regression vs multiple linear regression: A guide with examples. *Medium*. medium.com/@joshibhagyesh29/understanding-simple-linear-regression-vs-multiple-linear-regression-a-guide-with-examples-c3e8945e4830 (13 August 2025).

Kandel, A., & Pokhrel, K., 2024: Study of Urban Sprawl and Its Impact on Vegetation, Land Surface Temperature and Air Pollution Using Remote Sensing and GIS in Kathmandu Valley From 2015 to 2020. *Journal of Geoscience and Environment Protection*, 12(3), 28-53. doi.org/10.4236/gep.2024.123003.

Kaplan, G., Avdan, Z. Y., & Avdan, U., 2019: Spaceborne nitrogen dioxide observations from the Sentinel-5P TROPOMI over Turkey. *International Electronic Conference on Remote Sensing*, 18(1), 4. doi.org/10.3390/ECRS-3-06181.

Kenton, W., 2025. What is analysis of variance (ANOVA)? Investopedia, investopedia.com/terms/a/anova.asp (21 August 2025).

Khurshid, S., Loganathan, B., Duvinage, M., 2024: Comparative Evaluation of Applicability Domain Definition Methods for Regression Models. doi.org/10.48550/arXiv.2411.00920.

Luijken, K., Groenwold, R. H. H., Van Calster, B., Steyerberg, E. W., Van Smeden, M., 2019: Impact of predictor measurement heterogeneity across settings on the performance of prediction models: A measurement error perspective. *Statistics in Medicine,* 38(18), 3444–3459. doi.org/10.1002/sim.8183.

Manila Observatory., 2022. New Year's Eve 2022 particle pollution measurements in Metro Manila. observatory.ph/2022/03/29/new-years-eve-2022-particle-pollution-measurements-in-metro-manila (28 November 2024).

National Jewish Health., n.d.: Summer ozone dangers. nationaljewish.org/education/health-information/air-pollution-and-healthy-homes/outdoor-air-pollution/summer-ozone-dangers# (28 November 2024).

Nwagoum, C. S. K., Yemefack, M., Tedou, F. B. S., & Oben, F. T., 2023: Sentinel-2 and Landsat-8 potentials for high-resolution mapping of the shifting agricultural landscape mosaic systems of southern Cameroon. *International Journal of Applied Earth Observation and Geoinformation*, 124, 103545. doi.org/10.1016/j.jag.2023.103545.

Olfato-Parojinog, A., Sobremonte-Maglipon, P. A., Limbo-Dizon, J. E., Almadrones-Reyes, K. J., & Dagamac, N. H. A., 2023: Land use/land cover changes (LULCC) using remote sensing analyses in Rizal, Philippines. *GeoJournal*, 88(6), 6105–6118. doi.org/10.1007/s10708-023-10959-7.

Prana Air., 2024. What is temperature? *Health impacts.* pranaair.com/what-is-temperature (27 November 2024).

Portet S., 2020: A primer on model selection using the Akaike Information Criterion. *Infect Dis Model,* 5, 111-128. doi.org/10.1016/j.idm.2019.12.010.

Roces, I., 2023. Grab Philippines is prepared for the higher volume of the Holiday rush. Manila Bulletin. mb.com.ph/2023/12/4/grab-philippines-is-prepared-for-the-higher-volume-of-the-holiday-rush (2 December 2024).

Wells, B., 2025. Nitrogen Dioxide's Transformation: Pollutant formation process. ShunWaste. shunwaste.com/article/how-do-nitrogen-dioxide-become-a-pollutants-form (17 August 2025).

Yadav, A., 2023. Atmospheric Chemistry of Urban Pollutants and their Photochemical Transformations. *International Journal of Research and Analytical Review (IJRAR),* 10 (1), 195-200.

Yateswrecker., 2024. How to mitigate the impact of heat on fuel efficiency. *Doug Yates Towing & Recovery.* yateswrecker.com/blog/2024/06/27/the-impact-of-heat-on-fuel-efficiency (1 December 2024).