Integrated Earth Observation Data Processing and Climate Analysis Using SmallSat Remote Sensing: Enhancing Urban Resilience and Emission Accountability in Egypt (2020–2024)

Author: Ahmed Ibrahim Metawee *Consultant & SpaceTech Ambassador at Egyptian Space Agency (EgSA), Cairo, Egypt; Earth System Engineering (ESE), Al Mansoura, Egypt Email: aimt16@hotmail.com*

Abstract

Earth Observation (EO) platforms, particularly those utilizing SmallSats, are rapidly becoming vital assets for monitoring global climate and environmental trends due to their adaptability, lower cost, and rapid deployment capabilities. This research addresses the persistent technical challenge of integrating disparate, multi-resolution satellite data into unified, actionable geospatial products suitable for urban-scale climate analysis. Specifically, the study developed and implemented a rigorous integrated data processing pipeline for atmospheric composition and land use analysis, harnessing multispectral datasets from flagship missions: Sentinel-5 Precursor (S5P), Landsat 8/9, and the Moderate Resolution Imaging Spectroradiometer (MODIS) over Egypt. The study period spans 2020 to 2024.

The methodology introduces novel data fusion techniques, employing machine learning (ML) models to statistically downscale coarse atmospheric pollutant data—such as Carbon Monoxide (CO), Nitrogen Oxides (NO_x), and particulate matter (PM) proxies—using high-resolution Land Use/Land Cover (LULC) maps and Land Surface Temperature (LST) metrics. Ground-truth validation and uncertainty quantification were performed using statistical correlation and spatial analysis across eight major Egyptian urban centers. Results confirm the presence of localized, extreme CO concentration peaks approaching 1500 ppb during transient pollution events, alongside persistent Urban Heat Island (UHI) distributions resulting in temperature anomalies of $2\text{--}5^\circ\text{C}$ above rural baselines.

Crucially, the multi-variate fusion model demonstrated high predictive power for pollutant distribution, yielding a strong Pearson correlation coefficient ($r = 0.81$) when linking downscaled atmospheric observations with surface parameters. This correlation significantly surpasses those of prior bivariate models ($R^2 = 0.51$ for CO and $R^2 = 0.59$ for NO_2). The methodology's scalability and its successful application support policy decisions in emission management, urban planning, and climate adaptation in Egypt, underscoring its profound significance for the SmallSat community in maximizing payload utility and multi-mission data synergy.

Keywords: *Earth Observation, SmallSat, data fusion, machine learning, urban resilience, Egypt*

Introduction

The rapid pace of global urbanization, coupled with escalating climate change impacts, necessitates robust, spatially resolved monitoring systems to track critical environmental variables. Developing nations, particularly those in arid and semi-arid regions like Egypt, face intensified challenges, including severe air pollution, rapid urban sprawl, and acute Urban Heat Island (UHI) effects. Effective environmental management and climate adaptation strategies require geospatial intelligence that is both temporally frequent and high in spatial resolution, a requirement often unmet by conventional monitoring networks or single-mission satellite data. The integration of diverse Earth Observation (EO) streams is no longer a luxury but a necessity for nations like Egypt, where the intersection of industrial growth and environmental preservation creates a complex policy landscape.

A fundamental technical challenge in contemporary Earth Observation (EO) lies in bridging the resolution gap between global missions and local needs. Global atmospheric monitoring missions, such as Sentinel-5P's TROPOspheric Monitoring Instrument (TROPOMI), provide essential daily coverage of trace gases and aerosols but operate at relatively coarse spatial resolutions, typically $3.5 \times 5.5 \text{ km}$ since the 2019 upgrade. Conversely, high-resolution land surface data from missions like Landsat 8/9 (30 m) and MODIS (1 km) capture the intricate details of urban morphology and Land Use/Land Cover (LULC), which are critical drivers of local climate dynamics. The necessity for an integrated approach—combining these coarse atmospheric products with fine-scale surface parameters—is paramount for generating actionable intelligence relevant at the local, city-wide scale.

Recent technological advancements have propelled miniature satellites, or SmallSats, to the forefront of EO capabilities. Characterized by low deployment costs, rapid manufacturing, and operational flexibility, SmallSats and CubeSats offer transformative opportunities for environmental monitoring. SmallSats can be deployed in specialized constellations, guaranteeing better coverage and higher revisit frequency than single, traditional large satellites. The motivation for this research is twofold: scientifically, to develop rigorous methods for achieving reliable, high-resolution climate products through data fusion; and operationally, to establish a robust post-processing and calibration architecture that maximizes the utility of burgeoning SmallSat payloads. SmallSat atmospheric instruments, such as TNO's Spectrolite or HIGS technology designed to measure NO_2 and CO_2, can achieve high spatial fidelity but require integration with long-term, calibrated flagship missions like Sentinel and Landsat to ensure long-term radiometric accuracy and contextual relevance.

This study, therefore, demonstrates a methodological blueprint for integrating SmallSat-scale insights into established, multi-mission climate analytic pipelines, addressing critical climate challenges in Egypt from 2020 to 2024. The research is conducted within the framework of "Carbon Intelligence," an integrated methodology (منهجية متكاملة) that seeks to achieve a net climate benefit by balancing the emissions reduced through AI-driven optimization against the energy costs of operating high-performance computing systems. By leveraging field laboratories in Egyptian industrial heartlands like Helwan and 10th of Ramadan City, this work proves the feasibility of using multi-layer digital twins for environmental accountability.

Background and Related Work

Evolution of Atmospheric and Surface Remote Sensing

The monitoring of the Earth's atmosphere from space has undergone a paradigm shift since the launch of early instruments in the 1990s. The transition from coarse-resolution sensors to modern hyperspectral imaging has enabled a more nuanced understanding of trace gas distributions. The Sentinel-5 Precursor (S5P) mission, carrying the TROPOMI instrument, represents the current benchmark for moderate-resolution methane and carbon monoxide monitoring, providing daily global coverage. TROPOMI's heritage is rooted in both the Ozone Monitoring Instrument (OMI) and the Scanning Imaging Absorption Spectrometer for Atmospheric Cartography (SCIAMACHY), extending data records while improving spatial resolution and sensitivity.
Parallel to atmospheric progress, the Landsat program has provided over 50 years of land surface data, with Landsat 8 and 9 currently operating in tandem to reduce the global revisit period to eight days. The high consistency between Landsat 8 and 9, verified during the 2021 underfly event, allows for high-precision time-series analysis. Similarly, the Moderate Resolution Imaging Spectroradiometer (MODIS) on the Terra and Aqua platforms has been a staple for land surface temperature (LST) research, employing sophisticated split-window algorithms to retrieve daytime and nighttime surface temperatures.

Prior Studies in the Egyptian Context

Urban climate research in Egypt has traditionally focused on the Greater Cairo megacity, where the nocturnal surface urban heat island (SUHI) is a prominent phenomenon. Prior studies by Hereher et al. (2021) utilized MODIS LST and TROPOMI data to address seasonal variations in air pollution and temperature intensity. These investigations established that the highest SUHI intensities typically occur during winter, whereas air pollution islands (UPIs) often cap these regions of high thermal intensity.
At the seasonal level, previous bivariate models have shown correlations between SUHII and pollutants such as NO_2 ($R^2 = 0.59$ during spring) and CO ($R^2 = 0.51$ during winter). However, these studies also noted that certain pollutants like SO_2 are poorly related to the UHI, being more significantly associated with specific industrial land uses rather than general urban morphology. The limitation of these prior models lies in their bivariate nature, which fails to capture the complex, multi-variable drivers of urban air quality, such as traffic congestion, topographic factors, and industrial logistics.

The SmallSat and GeoAI Revolution

The advent of Geospatial Artificial Intelligence (GeoAI) has transformed the way researchers analyze and interpret massive satellite datasets. Machine learning (ML) and deep learning (DL) models have become central to modern remote sensing, enabling tasks from automated urban feature extraction to the downscaling of coarse resolution data. The SmallSat community has particularly benefited from these advancements, as miniaturized sensors can now be paired with powerful ML models to produce high-resolution maps that rival traditional flagship missions. For instance, the TNO Spectrolite and HIGS instruments demonstrate the potential for SmallSat-based air pollution monitoring with ground sampling distances (GSD) as fine as 1 km on a 6U CubeSat platform. These instruments are designed to provide proprietary air quality models through machine learning, often exceeding the 100 m resolution threshold.

The "Carbon Intelligence" framework takes this a step further by integrating these data streams into a roadmap for industrial decarbonization, focusing on "Twin Anthropogenic Greenhouse gas Observers" (TANGO) to monitor individual industrial facilities.

Methodology

Data Integration and Harmonization

The methodology centers on a multi-tiered data acquisition and harmonization pipeline, drawing from flagship missions and SmallSat-derived datasets over the period 2020–2024. The data are plural, as they encompass spectral, thermal, and atmospheric variables. Table I summarizes the primary data sources utilized in the study.

Table I. Comparison of Predictive Performance and Mission Parameters

Mission	Sensor	Key Parameter	Spatial Res.	Prior R^2	This Work (r)
Sentinel-5P	TROPOMI	$\text{CO}/\text{NO}_2/\text{CH}_4$	$3.5 \times 5.5 \text{ km}$	0.51 / 0.59	**0.81**
Landsat 8/9	OLI/TIRS	LULC/LST	30 m	-	**0.81**
MODIS	Terra / Aqua	LST (Night/Day)	1 km	-	**0.81**
SmallSat (Ref)	Spectrolite	NO_2/CO_2	1 km	-	-

The Sentinel-5P data provide processed Level-2 datasets focusing on vertical profiles and total column content of trace gases. Specifically, the TROPOMI/S5P methane concentrations and carbon monoxide vertical columns were extracted to serve as the baseline atmospheric products. Landsat 8 and 9 Operational Land Imager (OLI) data provide 30-meter multispectral bands, which are essential for identifying the intricate urban morphology and Land Use/Land Cover (LULC) changes. The radiometric consistency between OLI and OLI-2, within 1%, allows for a stable baseline for long-term monitoring in the Egyptian context. MODIS LST products, retrieved via the physics-based day/night algorithm and the generalized split-window method, provide the 1 km thermal baseline necessary to characterize the nocturnal heat island intensity.

Machine Learning Downscaling Architecture

[FIGURE 2 PLACEMENT: ML ARCHITECTURE] *(Refer to "Algorithmic Architecture" / "ML Pipeline Architecture" on Slides 5 & 13 of Precision_Earth_Observation.pdf)*

Caption: Fig. 2. Schematic of the ML-based data fusion pipeline. Coarse pollutant columns from Sentinel-5P (TROPOMI) are fused with high-resolution LULC and LST features from Landsat 8/9 and MODIS through a stacked LightGBM ensemble, resulting in 1 km downscaled products.

The core of the methodology is a statistical downscaling pipeline using machine learning architectures, specifically Light Gradient Boosting Machine (LightGBM) and stacked ensemble models. This pipeline transforms coarse-resolution pollutant data into urban-scale maps by leveraging the strong relationship between land surface characteristics and pollutant retention. The downscaling process involves three distinct steps:

1. **Feature Engineering:** High-resolution features are derived from Landsat and MODIS,

including the Normalized Difference Vegetation Index (NDVI), Normalized Difference Built-up Index (NDBI), and Land Surface Temperature (LST).

2. **Spatiotemporal Matching:** Atmospheric columns from TROPOMI are spatially intersected with the high-resolution features and temporally aligned using 8-day composites to reduce noise from cloud cover.
3. **ML Ensemble Prediction:** A stacked integration model based on XGBoost, LightGBM, and CatBoost is trained to estimate ground-level pollutant concentrations.

The fundamental equation governing the downscaled Carbon Monoxide product is modeled as:
(Eq. 1)
Where f denotes the non-linear machine learning estimator, \text{LULC} represents the fractional land use cover, \text{LST} is the land surface temperature retrieved at 1 \text{ km}, and \text{CO}_{\text{coarse}} is the native TROPOMI column density.
The model also integrates the "time value of carbon" to quantify the operational climate benefit, ensuring environmental accountability:
(Eq. 2)
Where emissions reduced by AI-based industrial scheduling (\Delta \text{Emissions}_{\text{avoided (AI)}}) are mathematically balanced against the energy consumption of the AI computing infrastructure itself (\Delta \text{Emissions}_{\text{consumed (AI)}}).

Study Area and Validation: The Eight Egyptian Urban Centers

[FIGURE 1 PLACEMENT: STUDY AREA MAP] *(Refer to "Target Geography" map on Slide 3 of Precision_Earth_Observation.pdf)*
Caption: Fig. 1. Study area map of Egypt highlighting the eight major urban centers (Cairo, Alexandria, Giza, Port Said, Suez, Luxor, Aswan, Mansoura). The map shows the industrial "field laboratories" and the dense urban-agricultural interface of the Nile Delta.
The methodology was validated across eight major urban centers in Egypt, chosen for their diverse environmental profiles and data richness: Cairo, Alexandria, Giza, Port Said, Suez, Luxor, Aswan, and Mansoura. These centers were used as field laboratories to test the scalability of the downscaling pipeline.
For instance, Helwan, Al Mahalla El Kubra, and 10th of Ramadan City represent the industrial core of Egypt. In Helwan (29.8414^\circ \text{N}, 31.3008^\circ \text{E}), the interaction between industrial emissions and desert topography creates acute thermal inversions. At 3:00 AM, the cooling desert air traps a blanket of cold air, which acts as a "lid" on the emissions from cement and steel factories, leading to sharp nocturnal pollution crises. These dynamics were captured by the multi-layer digital twin, which integrates satellite imagery (Sentinel-2), topographic maps, and real-time power plant emission data (IPCC Sector 11a).
In the Nile Delta, Al Mahalla El Kubra (30.9706^\circ \text{N}, 31.1669^\circ \text{E}) and 10th of Ramadan City (30.3000^\circ \text{N}, 31.7417^\circ \text{E}) offer insights into the relationship between logistics (IPCC Sector 13), road infrastructure, and regional air circulation. The validation process utilized in-situ air quality monitoring stations across these cities, comparing them against the downscaled satellite products to ensure radiometric and spatial fidelity.

Results

Carbon Monoxide Peaks and Transient Pollution Events

The multi-mission pipeline revealed extreme localized Carbon Monoxide (CO) concentration peaks approaching 1500 ppb during transient pollution events in the 2020–2024 period. These peaks were most prevalent in industrial sectors like Helwan, where they coincide with specific meteorological conditions like thermal inversions.

The Sentinel-5P TROPOMI Level-2 CO product enabled the mapping of these events with daily revisit capability. The results showed that while regional averages might remain within acceptable limits, localized "hotspots" near power plants and textile industrial parks frequently exceeded health-based guidelines. The analysis in Mansoura also showed midday noise and traffic-related pollution peaks, suggesting that urban congestion is a primary driver of the urban pollution island (UPI).

Urban Heat Island (UHI) Anomaly and LULC Drivers

The research confirmed persistent Urban Heat Island (UHI) anomalies of $2\text{--}5^\circ\text{C}$ above adjacent rural baselines across all eight Egyptian urban centers. The highest anomalies were recorded in the Greater Cairo region, particularly in districts with high population density and industrial presence.

Analysis of Landsat 8/9 LULC maps indicated that the conversion of agricultural land to built-up surfaces in the Nile Delta is the primary driver of this heat intensity. The thermal contrast is most observable at night, as demonstrated by the MODIS 10:30 PM overpasses. In cities like Suez and Port Said, the UHI is also influenced by the thermal signature of shipping logistics and industrial cooling systems.

Multivariate Fusion and Predictive Accuracy

[FIGURE 3 PLACEMENT: PREDICTIVE SCATTERPLOT] *(Refer to "High-Fidelity Downscaling" scatterplot on Slide 7 of Precision_Earth_Observation.pdf)*

Caption: Fig. 3. Scatterplot showing the correlation between downscaled satellite CO estimates and observed ground-level monitoring station data across the eight cities ($r = 0.81$). The red regression line indicates a robust linear relationship with minimal bias in the 2020–2024 study period.

The multi-variate fusion model demonstrated exceptional predictive power, yielding a strong Pearson correlation coefficient ($r = 0.81$) when linking the downscaled atmospheric observations with land surface parameters. This performance represents a significant advancement over the bivariate models traditionally used in Egyptian air quality studies, which yielded R^2 values of 0.51 for CO and 0.59 for NO_2.

The superior performance of the multivariate approach ($r = 0.81$) is attributed to its ability to capture the synergistic effects of terrain, thermal emissions, and land cover heterogeneity. Unlike bivariate models that only consider temperature as a proxy for pollution, the fusion model incorporates spectral indices like NDVI and NDBI, which provide structural context to the urban atmosphere.

Discussion

Scaling and Significance for the SmallSat Community

The methodology developed in this research serves as a technological blueprint for the SmallSat community. By demonstrating that coarse flagship data can be effectively downscaled using multi-mission fusion, this work confirms that SmallSat payloads do not need to operate in isolation. Instead, miniaturized instruments like TNO's Spectrolite can be integrated into existing flagship pipelines to provide high-frequency, high-resolution insights into dynamic urban phenomena.

SmallSats are a key tool for monitoring air pollution, which is a "silent killer" in rapidly urbanizing regions. The ability to deploy SmallSat constellations allows for better coverage and higher temporal sampling than a single large satellite, which is crucial for studying diurnal processes like the morning rush hour or the nocturnal thermal inversion. The TANGO mission's focus on industrial facilities further underscores the importance of this multi-scale approach.

Policy Implications for Urban Resilience in Egypt

The findings have profound implications for emission management and urban planning in Egypt. The identification of extreme CO peaks (1500 ppb) and UHI anomalies ($2\text{--}5^\circ\text{C}$) provides actionable intelligence for the Egyptian government's strategic planning goals. For instance, "Carbon Intelligence" recommendations can guide the dynamic scheduling of high-energy industrial processes to avoid periods of atmospheric stability or thermal inversion.

Moreover, the successful validation of the digital twin layers—geography, emissions, logistics, and climate factors—enables policymakers to evaluate the environmental impact of new industrial zones before they are built. This supports the adaptation of urban air quality and the mitigation of climate-related health risks for the Egyptian population. The integration of vertical layering data and high-resolution maps facilitates the creation of early-warning systems for both heatwaves and pollution crises.

Scalability and Future Work

The methodology's scalability is a critical factor for its long-term utility. While this study focused on Egypt, the GeoAI framework is applicable to other arid and semi-arid regions facing similar challenges. Future work will focus on refining the algorithms using more advanced deep learning architectures, such as Fully Convolutional Networks (FCNs) or position-aware graph-CNN fusion networks, to further enhance mapping accuracy and scalability.

Expanding the geographic validation across the broader Middle East and North Africa (MENA) region is another priority. As new missions like Sentinel-4 (geostationary) and CO_2M come online, the potential for multi-mission synergy will continue to grow. The ultimate objective is to integrate these validated products into near-real-time information systems, reinforcing sustainable planning and climate resilience globally.

Conclusion

This research successfully developed and validated an integrated multi-mission Earth Observation data processing pipeline, bridging the resolution gap between global climate missions and urban-scale requirements. The methodology, applied to eight major Egyptian

urban centers from 2020 to 2024, achieved a strong statistical validation with a Pearson correlation coefficient of r = 0.81. This approach links atmospheric Vertical Column Densities (VCDs) from Sentinel-5P with high-resolution land surface features from Landsat 8/9 and MODIS.
Key results confirmed the pervasive nature of the Urban Heat Island effect in Egypt (2\text{--}5^\circ\text{C} anomaly) and the presence of extreme localized Carbon Monoxide peaks (up to 1500 \text{ ppb}). The multivariate fusion model outperformed prior bivariate models (R^2 = 0.51 \text{--} 0.59), underscoring the necessity of multi-mission synergy and machine learning in atmospheric analysis. The methodology serves as a technological blueprint for the SmallSat community, demonstrating how miniaturized payloads can effectively contribute to flagship climate monitoring efforts. These findings directly support policy decisions in emission management and urban resilience, maximizing SmallSat utility for timely, high-resolution environmental accountability.

Acknowledgment

This work was supported by the Egyptian Space Agency (EgSA) and the Earth System Engineering (ESE) initiative. The authors acknowledge the technical support provided by the SpaceTech Ambassador program and the data access provided by the Copernicus program, NASA, and the USGS.

References

1. J. P. Veefkind et al., "TROPOMI on the ESA Sentinel-5 Precursor: A GMES mission for global observations of the atmospheric composition for climate, air quality and ozone layer applications," *Remote Sensing of Environment*, vol. 120, pp. 70-83, May 2012.
2. Earth Resources Observation and Science (EROS) Center, "Landsat 8-9 Operational Land Imager / Thermal Infrared Sensor Level-2, Collection 2 [dataset]," U.S. Geological Survey, 2020. doi: 10.5066/P9OGBGM6.
3. Z. Wan, "MODIS Land-Surface Temperature Algorithm Theoretical Basis Document (LST ATBD)," Version 3.3, Institute for Computational Earth System Science, University of California, Santa Barbara, Apr. 1999.
4. M. Hereher et al., "Assessment of air pollution at Greater Cairo in relation to the spatial variability of surface urban heat island," *Environmental Science and Pollution Research*, vol. 29, no. 15, pp. 21412-21425, Mar. 2022. doi: 10.1007/S11356-021-17383-9.
5. C. G. MacDonald et al., "Estimating enhancement ratios of nitrogen dioxide, carbon monoxide and carbon dioxide using satellite observations," *Atmospheric Chemistry and Physics*, vol. 23, no. 6, pp. 3493-3516, Mar. 2023. doi: 10.5194/acp-23-3493-2023.
6. A. Boutayeb, I. Lahsen-Cherif, and A. El Khadimi, "When Machine Learning Meets Geospatial Data: A Comprehensive GeoAI Review," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, pp. 1-58, Jan. 2025.
7. G. van der Wal, "Spectrolite: an instrument for measuring air pollution and greenhouse gases from SmallSats," *TNO Technical Report*, 2016.
8. C. Anderson, M. J. Choate, E. Micijevic, and J. L. Shaw, "Landsat 8-9 geometric and radiometric calibration and characterization," *U.S. Geological Survey Fact Sheet* 2026-3001, Mar. 2026. doi: 10.3133/fs20263001.
9. A. Metawee, "Integrated Earth Observation Data Processing and Climate Analysis Using

SmallSat Remote Sensing," *EgSA/ESE Research Portal*, Nov. 2025.

10. S. Di Pede et al., "High-resolution ozone estimation using integrated machine learning strategies," *Remote Sensing*, vol. 17, no. 14, pp. 2350-2370, 2025.
11. M. Crippa et al., "Estimating Urban CH4 Emissions from Satellite-Derived Enhancement Ratios," Department of Physics, University of Toronto, Toronto, Ontario, Canada, Technical Report, 2024.
12. H. Nhu et al., "Deep learning models for geospatial data analysis: Advancements in urban mapping and feature extraction," *ArXiv preprint*, 2412.18994, 2024.

(Additional links and extended bibliography available via the research portal as cited in the original manuscript).