Linking Socioeconomic Status and Emissions: The Predictive Power of the International Wealth Index for NO2 column densities

Francesco Grieco
fra.grieco@live.com

Chalmers University of Technology
Adel Daoud
Chalmers University of Technology
Mohammad Kakooei
Chalmers University of Technology

Article

Keywords: International Wealth Index (IWI), NO2 monitoring, Machine learning in environmental policy, Nationally Determined Contributions (NDCs)

Posted Date: April 8th, 2025

DOI: https://doi.org/10.21203/rs.3.rs-6347584/v1

Additional Declarations: There is **NO** Competing Interest.

Linking Socioeconomic Status and Emissions: The Predictive Power of the International Wealth Index for NO_2 column densities

Francesco Grieco[1,2*], Mohammad Kakooei[1,2] and Adel Daoud[1,2]

[1*]Department of Computer Science, Chalmers University of Technology, Gothenburg, Sweden.
[2]AI and Global Development Laboratory, Linköping University, Linköping, Sweden.

*Corresponding author(s). E-mail(s): francesco.grieco@chalmers.se;
Contributing authors: kakooei@chalmers.se; daoud@chalmers.se;

Abstract

Ensuring accountability in emissions reductions is critical as many nations struggle to meet their Nationally Determined Contributions (NDCs) under the Paris Agreement. Traditional greenhouse gas (GHG) monitoring approaches are hindered by the long atmospheric lifetimes of key gases like carbon dioxide (CO_2) and methane (CH_4). In contrast, nitrogen dioxide (NO_2) is a short-lived pollutant with high spatial and temporal variability, making it an effective proxy for tracking anthropogenic emissions. This study introduces a novel predictive framework integrating satellite-derived NO_2 data with socioeconomic indicators, specifically the International Wealth Index (IWI). Using machine learning techniques, we establish IWI as a reliable predictor of NO_2 column densities, demonstrating that socioeconomic development patterns significantly influence emission trends. Our convolutional neural network (CNN) model achieves an average predictive accuracy ($R^2 = 0.56$) across the African continent, allowing for anticipatory analysis of emission hotspots. Forecasts for 2030 reveal substantial disparities between projected NO_2 levels and NDC commitments, highlighting regions at risk of non-compliance. These findings emphasize the potential of integrating socioeconomic and environmental data to improve emissions monitoring, inform policy decisions, and enhance accountability in global climate action.

Keywords: International Wealth Index (IWI), NO_2 monitoring, Machine learning in environmental policy, Nationally Determined Contributions (NDCs)

Main

When future emissions projections based on Nationally Determined Contributions (NDCs) fall short of the Paris Agreement's goals[1], and many nations struggle to meet even these pledges[2], stronger accountability becomes essential. Ensuring that states uphold their commitments is crucial to bridging this gap and advancing global climate action. Despite the global commitments established under the Paris Agreement and reinforced through subsequent COP meetings, there remains no robust mechanism to systematically monitor and enforce these promises[3]. Traditional methods of tracking greenhouse gas (GHG) emissions are limited by the long atmospheric lifetimes and complex vertical distributions of key gases like carbon dioxide (CO_2) and methane (CH_4), which hinder the identification of emission sources and trends[4]. CO_2 persists for decades to millennia[5], while CH_4 remains in the atmosphere for years[6], allowing these gases to mix across multiple atmospheric layers.
In contrast, NO_2 is a short-lived pollutant, with a lifetime of just a few hours during daytime, predominantly concentrated near the ground. This unique property makes NO_2 an exceptional proxy for tracking anthropogenic emissions[7]. Its strong temporal and spatial variability reflects emission patterns from localized sources, such as transportation networks, industrial activities, and energy production sites[8]. Furthermore, NO_2 is often co-emitted with other GHGs, making it a valuable and comprehensive indicator of human impact on the environment[9]. In addition to its utility as a proxy, NO_2 is independently significant as an air pollutant due to its harmful effects on public health. Exposure to elevated NO_2 levels has been linked to respiratory illnesses, cardiovascular problems, and premature mortality[10].

The advent of satellite-based observations, such as those from the Tropospheric Monitoring Instrument (TROPOMI), has revolutionized NO_2 monitoring. TROPOMI provides high-resolution global measurements, capable of pinpointing emission hotspots at the urban scale[11]. This unprecedented spatial detail, together with the atmospheric lifetime and distribution of NO_2, enables both the evaluation of compliance with NDC targets and the identification of local emission sources within individual countries (i.e. at the intra-country level).

However, translating these observational data into actionable insights for policymakers remains a significant challenge.

This study seeks to address this gap by integrating socioeconomic indicators with satellite-based NO_2 observations. Specifically, this article is the first to link the International Wealth Index (IWI) as a predictor of NO_2 column densities. The IWI measures critical dimensions of socioeconomic development, including material wealth distribution, which in turn is associated with urbanization and infrastructure growth, factors well-known to affect emission patterns[12]. As delineated in the Methods section, we use a novel IWI dataset originally covering Africa from 1990 to 2020 at a 6.7 km resolution[13]. To align with TROPOMI's full-year data availability since 2019, we applied their model to generate annual IWI estimates for 2019–2024. This study links these data with TROPOMI observations over the same period, leveraging machine learning (ML) techniques to establish a predictive framework that links socioeconomic data and environmental monitoring.

The predictive framework developed here is built on the hypothesis that higher levels of and changes in IWI can serve as a reliable indicator of future NO_2 column densities. Projections of IWI allow for anticipatory analysis, enabling the identification of regions where column densities are likely to increase, stagnate, or decrease. This approach moves beyond descriptive analysis, offering predictive insights that can guide targeted policy interventions and resource allocation.

Furthermore, our framework facilitates the evaluation of progress toward NDC goals. By forecasting future NO_2 column densities and comparing them to NDC targets, this study highlights regions and nations at risk of non-compliance. It provides an evidence base for policymakers to refine strategies, prioritize actions, and allocate resources effectively. Additionally, the high-resolution spatial data allows for the observation of intra-country column densities variations, offering insights into localized sources of emissions and enabling targeted interventions at the subnational level.

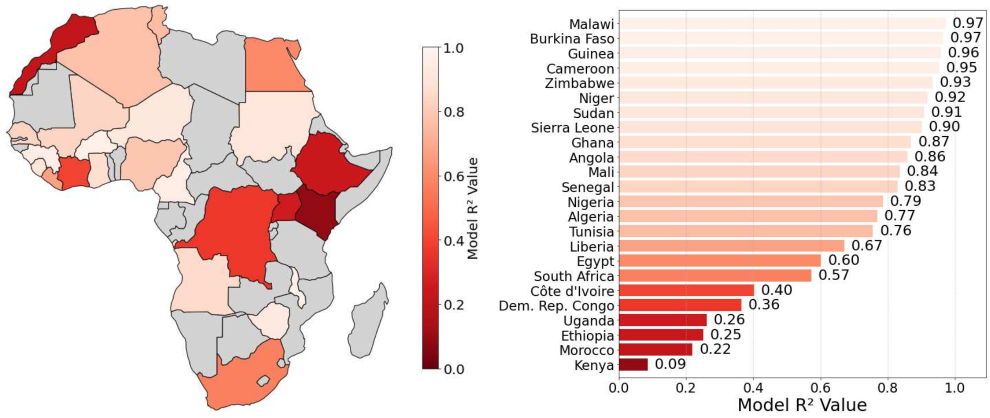
In summary, this study combines satellite-derived NO_2 data with socioeconomic projections to address key gaps in emissions monitoring and policy evaluation. The use of NO_2 as a short-lived proxy for GHG emissions, coupled with the predictive power of the IWI, offers a novel and scalable methodology for assessing progress toward global climate commitments. This study focuses on Africa as a case study due to the reliable and extensive IWI predictions available for this region from [13], providing a consistent dataset to validate our methodology. The focus is not meant to disproportionately scrutinize less culpable nations for climate change but to leverage high-quality socioeconomic data to showcase the broader applicability of our approach. This integrative approach not only informs climate action but also addresses broader sustainability goals, including public health and equity. This study contributes to the broader effort to hold nations accountable for their commitments under the Paris Agreement while, at the same time, identifying emission sources and enabling targeted interventions at the subnational level.

Results

Predictive power of IWI on NO_2

Our model, a 3-layer convolutional neural network (CNN), demonstrated an average predictive accuracy, measured by R^2, of 0.56 for predicting NO_2 column densities from IWI. Fig.2 provides a schematic representation of our model's framework. The model's performance surpasses that of more complex architectures, including deeper CNNs, VGG16-based models, and CNNs incorporating wind data as an additional input. Moreover, it significantly outperformed traditional machine learning approaches, such as linear regression and random forest regressors, highlighting its effectiveness and efficiency for this task (see Methods).

Figure 1b illustrates the predictive performance of the model across different African countries, represented by the R^2 value obtained when each country is used as a geographically isolated test dataset while training is performed on the rest of



(a) (b)

Fig. 1: Model performance (R^2) across African countries. R^2 values are obtained by training the model on the rest of the continent and using each country as a geographically isolated test dataset. Lighter shades indicate higher R^2 values, while gray regions represent countries with insufficient or missing data.

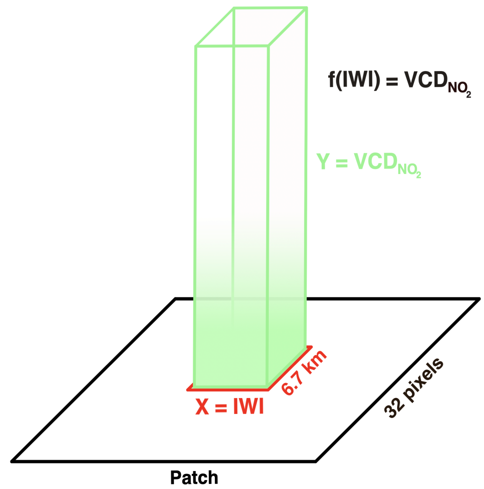


Fig. 2: Schematic representation of our model's framework.

the continent. The color gradient reflects the model's R^2, with lighter shades indicating higher predictive accuracy. Gray regions denote countries for which data were insufficient or not included in the analysis.

Lower R^2 values (darker red) are observed in countries such as Kenya, Ethiopia, and the Democratic Republic of Congo, suggesting that the model struggles to generalize well in these regions. This may result from unique socioeconomic or emission patterns that are not well-represented in the training data. Conversely, higher R^2 values (lighter shades) are observed in countries such as Mauritania, Mali, Niger, and Namibia, as well as in other parts of West Africa and the Sahara region. This better performance likely reflects more consistent emission patterns or better representation of these regions in the training data.

Intra-country forecasted changes in NO_2 column densities

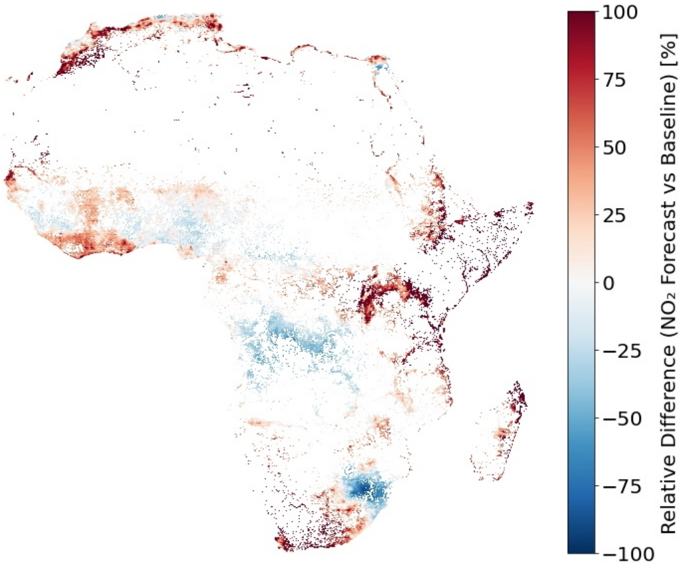


Fig. 3: Change in NO_2 column densities between 2019 and the forecasted 2030, over the populated areas of the African continent.

Figure 3 illustrates the differences in NO_2 column densities between 2030 predictions and 2019 observations from TROPOMI. To generate these predictions, we first obtain IWI estimates from an Earth Observation–Machine Learning (EO-ML) method, which provides gridded socioeconomic data at a 6.7 km resolution for Africa from 1990 to 2024[13]. We then extend this dataset by linearly interpolating IWI values forward to 2030. These projected IWI values are fed into our trained model, which predicts NO_2 column densities at the same spatial resolution. The model's output for 2030 is then compared to 2019 NO_2 column densities observed by TROPOMI, allowing us to quantify changes over time and identify regions experiencing notable increases or decreases in emissions. The map highlights both increases and decreases in column densities across the African continent, with notable intra-country variability. Significant intra-country variations in predicted NO_2 column densities are observed across the African continent. Hotspots of increased column densities are prominently observed in rapidly urbanizing and industrializing regions, such as East Africa (notably Kenya and

Ethiopia), Central Africa, and parts of West Africa. These increases are likely driven by urban expansion, industrial growth, and the associated rise in transportation activities. For instance, Kenya's industrialization efforts and Ethiopia's rapid urbanization align with these predictions, as seen in their urban development policies and economic growth targets[14, 15]. Additionally, such regions with limited implementation of air quality control measures are particularly susceptible to rising emissions[16]. Decreases in predicted NO_2 column densities are concentrated in parts of Southern Africa and localized areas in the Sahel region. These decreases may be attributed to a combination of socioeconomic transitions, shifts in energy sources, and the implementation of emission control policies[17]. For example, South Africa's Nationally Determined Contribution (NDC) commits to transitioning away from coal-fired power plants and adopting renewable energy sources, likely contributing to reduced emissions[18]. In the Sahel, sustainable agriculture and clean energy adoption have been gaining traction, reducing reliance on traditional biomass fuels and mitigating emissions from rural areas[19].

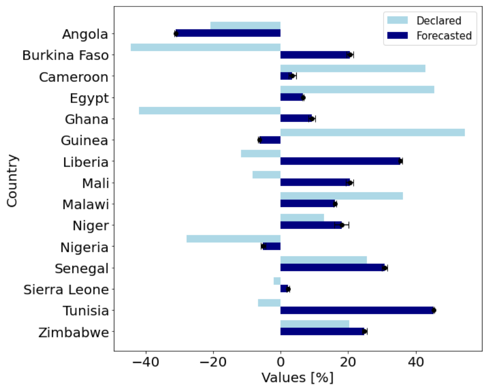
Disparities between forecasted NO_2 column densities and NDC goals

The results highlight substantial discrepancies between the forecasted difference in NO_2 column densities for 2030 and the relative reduction targets outlined in the NDCs for many African countries. Fig. 4a presents a histogram comparing the column density variation from 2019 to 2030 as predicted by the ML model versus the NDC goals for each country.

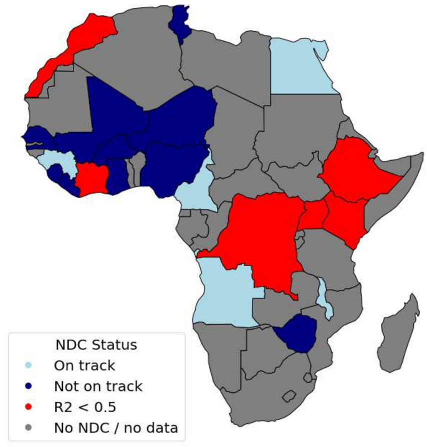
Some countries, such as Burkina Faso, Ghana, Liberia, Mali, Nigeria and Tunisia demonstrate substantial underperformance relative to their NDC targets, with forecasted column densities far exceeding their declared goals. These findings suggest significant challenges in meeting emission reduction commitments, potentially due to economic pressures or insufficient policy enforcement. On the other hand, countries like Cameroon, Egypt, Guinea and Malawi are on track to exceed their NDC targets, indicating either effective policy implementation or slower economic growth trajectories that naturally reduce emissions.

A geographical visualization of NDC compliance likelihood, shown in Fig. 4b, further illustrates these trends. Nations in West Africa, such as Burkina Faso and Nigeria, are at the highest risk of non-compliance, reflecting potential gaps in institutional capacity or infrastructure for implementing emission reduction strategies. Southern African countries like Zimbabwe display smaller gaps, indicating better alignment with NDC targets, though challenges remain.

These findings carry significant implications for global emission reduction efforts. Regions with the largest gaps between forecasted and declared emissions represent high-impact areas where targeted interventions could yield the greatest benefits. This highlights the importance of data-driven policy planning and international collaboration to address the capacity challenges faced by underperforming nations.



(a)



(b)

Fig. 4: (a) Histogram comparing NDC goals and predicted NO_2 column densitie variation between 2019 and 2030 by country. The countries shown are the ones for which the model's R^2 is higher than 0.5, when having the country as test dataset, and the ones for which NDC are available and for which there were enough pixels to calculate R^2. (b) Map of African countries showing likelihood of NDC compliance by 2030. Countries in red are those where the model's R^2 falls below 0.5 when the country is used as the test dataset. Countries in grey either have an R^2 lower than 0.5 under the same condition, lack available NDC data, or do not have a sufficient number of pixels to calculate R^2.

Discussion

This study demonstrates the predictive power of socioeconomic development, measured as the International Wealth Index (IWI), in forecasting NO_2 column densities. Our study highlights its potential in assessing progress toward Nationally Determined Contributions (NDCs) under the Paris Agreement. By integrating socioeconomic indicators with satellite-derived NO_2 data, we provide a novel framework for linking development patterns to emissions and identifying areas where emission control policies are most urgently needed.

The simplicity of the 3-layer CNN ensures computational efficiency while maintaining robust predictive performance, making it an accessible tool for large-scale analyses.

The predictive strength of IWI validates its utility as a proxy for socioeconomic factors driving emissions, such as urbanization, infrastructure development, and economic activity. However, it is important to note that predictive accuracy does not equate to causation. While our findings highlight the value of IWI for identifying column density patterns, future studies should aim to establish causal mechanisms linking specific socioeconomic activities to NO_2 emissions. For instance, differentiating emissions originating from transportation, industrial processes, and residential energy use could provide more actionable insights for policymakers.

Intra-country and regional insights

The spatially resolved predictions from our model reveal significant intra-country variability in NO_2 column densities. Hotspots of increased column densities are prominently observed in rapidly urbanizing and industrializing regions, such as Kenya and Ethiopia in East Africa, as well as parts of Central and West Africa. These findings align with documented trends in urban expansion and industrial growth, which drive transportation and energy demands[14, 15].

Conversely, areas of predicted column density decreases can be found in Southern Africa and localized regions within the Sahel. These reductions may reflect shifts in socioeconomic activities, transitions to cleaner energy sources, and the implementation of emission control policies[17, 19]. For example, South Africa's commitment to phasing out coal-fired power plants and the adoption of sustainable agriculture in the Sahel region have likely contributed to these trends[18, 16].

Our forecasting how socioeconomic changes will influence column densities at the local level is crucial for designing targeted interventions, particularly in regions with limited air quality control measures or institutional capacity.

Implications for NDC compliance

The comparison of forecasted difference in NO_2 column densities for 2030 with NDC targets reveals significant gaps for many African nations. Some countries, such as Burkina Faso and Nigeria, exhibit forecasted increases in NO_2 column densities that far exceed their stated reduction goals. This underscores the challenges posed by rapid urbanization, limited institutional capacity, and economic constraints in meeting climate commitments. In contrast, countries like Cameroon, Egypt, Guinea and Malawi are on track to exceed their NDC targets, potentially reflecting effective policy implementation or slower economic growth trajectories.

These disparities highlight the importance of integrating emissions monitoring with socioeconomic data to support data-driven policy planning. Our findings suggest that regions with the largest gaps between forecasted and declared emissions should be prioritized for capacity-building initiatives and international support.

Advancing emissions monitoring

The integration of IWI with satellite-derived NO_2 column densities represents a significant advancement in emissions monitoring. Unlike traditional methods that rely on self-reported inventories, our approach provides an independent, high-resolution assessment of emission trends. By leveraging the short atmospheric lifetime and ground-level presence of NO_2, our framework enables the identification of emission sources and hotspots with unprecedented spatial detail.

Moreover, the ability to observe intra-country variations in column densities offers new opportunities for subnational policy interventions. For instance, targeting specific urban centers or industrial zones identified as emission hotspots can maximize the impact of mitigation strategies. Additionally, our model provides a flexible framework for scenario-based analyses, allowing policymakers to assess the implications of various socioeconomic trajectories on future emissions.

Limitations and future directions

While this study provides valuable insights, several limitations should be acknowledged. The reliance on predictive models inherently introduces uncertainties, particularly in regions with sparse data or unique emission patterns. For example, countries with lower model performance, such as Kenya and Ethiopia, may require additional data or context-specific adjustments to improve predictive accuracy.

The linear interpolation of IWI data to forecast future column densities is another limitation. While this approach provides a baseline estimate, it does not account for nonlinear socioeconomic changes or unexpected policy shifts. Future research could explore more sophisticated methods for projecting IWI, such as time-series models, Long Short-Term Memory (LSTM), or transformer models.

Moreover, the generalizability of our findings beyond Africa remains an open question. Expanding this approach to other regions with reliable IWI predictions could validate its broader applicability and inform global efforts to monitor and reduce emissions.

Finally, NO_2 can also serve as a valuable proxy for greenhouse gas emissions, particularly CO_2. Since NO_2 is co-emitted with CO_2 in combustion-related activities, including transportation, industry, and power generation[9], future research could build on this study's framework to develop predictive models that estimate CO_2 emissions based on NO_2 satellite observations and predictions. Expanding this predictive framework to CO_2 would enhance its utility for climate monitoring and mitigation efforts, reinforcing the role of satellite-based observations in tracking anthropogenic emissions at multiple scales.

Methods

Data Description

International Wealth Index (IWI)

The International Wealth Index (IWI) is an asset-based measure of material well-being designed for comparability across low- and middle-income countries. Unlike other wealth indices that are often country-specific and context-bound, IWI offers a standardized approach for evaluating household wealth and socioeconomic status, allowing for cross-national and temporal comparisons. Developed using data from 2.1 million households across 97 countries, IWI is grounded in household asset ownership, access to basic services, and housing quality, effectively capturing material well-being on a global scale.

IWI evaluates a household's material standard of living using a carefully selected set of assets, including consumer durables like televisions, refrigerators, and bicycles; housing characteristics such as the quality of flooring and toilet facilities; and access to essential utilities like electricity and clean water. The index is constructed using Principal Component Analysis (PCA), which assigns weights to each asset based on its contribution to overall wealth. These weights are combined into a unified scale that runs from 0 to 100, where a score of 0 represents a household with no durable goods, poor-quality housing, and no access to utilities, and 100 represents a household possessing all assets and services included in the index[12].

The IWI is available only at the locations and times where household surveys have been conducted. To overcome this limitation, [13] developed an EO-ML method to infer neighborhood-level material-asset wealth across the populated areas of the African continent from 1990 to 2019. This method combines multi-temporal satellite imagery, available DHS surveys, and recurrent convolutional neural networks to produce spatially and temporally consistent estimates of IWI. For this study, we applied their model to obtain data for more recent years, generating annual IWI data for 2019–2024 to align with the operational period of the TROPOMI instrument. The resulting data is organized into gridded images at a resolution of 6.7 km.

This study focuses on Africa as a case study because [13] provide reliable IWI predictions for this region, offering a geographically extensive and temporally consistent dataset to validate our methodology. It is important to emphasize that the focus on the African continent is not intended to disproportionately scrutinize countries that are less responsible for climate change. Instead, it leverages the availability of high-quality socioeconomic data to demonstrate the broader applicability of our approach in regions with reliable information.

NO_2 data

Nitrogen dioxide (NO_2) is mainly generated by the combustion of fossil fuels, such as those used in vehicles, power plants, and industrial activities. Additionally, smaller contributions arise from natural sources like soil microbial activity and vegetation. Beyond its role as a direct pollutant, NO_2 acts as a precursor to other harmful atmospheric compounds, including ground-level ozone, fine particulate matter, and acid rain. Despite its significant health and environmental impacts, NO_2 is not categorized as a major greenhouse gas due to its relatively low global column densities—approximately three orders of magnitude lower than carbon dioxide (CO_2)—and its short atmospheric lifespan of a few hours to a day. This brevity ensures that NO_2 remains concentrated near emission sources, making it a valuable indicator for identifying pollution hotspots[20].

This study employs NO_2 measurements from the TROPOspheric Monitoring Instrument (TROPOMI), a high-resolution spectrometer aboard the Sentinel-5P satellite, launched in 2017. TROPOMI captures NO_2 column densities with a spatial resolution as fine as 5.5 km $\times$ 3.5 km, enabling the detection of medium-scale pollution sources, including urban centers, industrial facilities, highways and large individual emission events such as wildfires or shipping activities. By maintaining a fixed observation time of approximately 13:30 local solar time, TROPOMI ensures consistent measurements, reducing variability caused by daily fluctuations in NO_2 levels[21]. Extensive validation studies have confirmed the instrument's accuracy and reliability in monitoring atmospheric NO_2[22].

The instrument detects sunlight reflected from the Earth's surface and atmosphere, analyzing a broad range of wavelengths, from ultraviolet to shortwave infrared. Through these measurements, TROPOMI identifies absorption features specific to NO_2, allowing it to estimate the tropospheric vertical column density (VCD) of the gas[23]. This quantity serves as a proxy for ground-level emissions. In fact, although minor amounts of NO_2 can be transported from the stratosphere via processes such as photolysis of nitrous oxide (N_2O) and subsequent stratospheric intrusions, their impact on tropospheric column densities is negligible compared to surface-level sources[24].

For this analysis, we used yearly median of NO_2 data from 2019, the first complete year of TROPOMI observations, to 2024, re-gridded to match the resolution of the IWI data.

NDC data harmonization

Nationally Determined Contributions (NDCs) are central to the Paris Agreement's objective of limiting global temperature rise to well below 2°C above pre-industrial levels, with efforts to cap the increase at 1.5°C. Each signatory nation is mandated to formulate, communicate, and maintain successive NDCs, detailing their strategies for reducing greenhouse gas emissions and adapting to climate change impacts. NDCs embody the global commitment to combat climate change, with each country contributing according to its capabilities and responsibilities. By publicly outlining their climate actions, countries provide a basis for tracking progress[25].

Countries publish their emission reduction targets for various GHGs, summarized as CO_2 equivalents (CO_2e), for different target years and relative to different base years. This variability complicates direct comparisons between countries. Additionally, NDCs lack standardization: they are presented as separate documents with varying formats. Some NDCs include detailed sector-specific breakdowns and distinctions between conditional (dependent on external support, such as funding or technology transfer) and unconditional (self-financed or guaranteed) commitments. Others only specify an aggregate reduction target without elaborating on the methods or conditions for achieving it[26].

To address these inconsistencies, we standardized all NDCs by converting them to a common baseline year (2019, the first full year of TROPOMI observations used in this study) and a common forecast year (2030). This involved linearly interpolating or extrapolating data from the reference years provided in the NDCs. We then calculated the difference in total CO_2 equivalent emissions between 2019 and 2030 for each country. Where available, we prioritized conditional targets to reflect the maximum commitments pledged. However, countries that report their reductions only relative to a Business-As-Usual (BAU) scenario for their estimation year could not be directly compared within this framework, resulting in their exclusion from the analysis.

It is important to interpret these results cautiously, as the standardization process introduces uncertainties. The assumption of linear emission trajectories between base years and target years is a simplification and may not reflect the complex, sectoral-specific dynamics of emission reduction efforts.

Model architecture and training

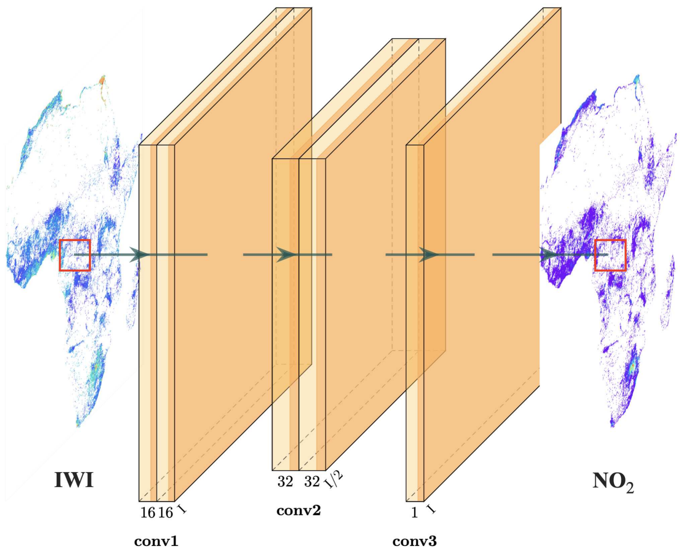


Fig. 5: Schematic representation of the CNN model used to predict NO_2 concentrations from IWI data. The model processes gridded IWI input through multiple convolutional layers (conv1–conv3) to extract spatial features and generate NO_2 predictions at the same resolution. The numbers indicate the number of filters in each convolutional layer, while Roman numerals denote the corresponding downsampling factors.

The primary model used in this study is a 3-layer convolutional neural network (CNN) (see Fig. 5). The model takes patches of size 32×32 pixels, each containing a single channel representing IWI values, as input, and corresponding NO_2 patches as targets. Patches were created by dividing the IWI and NO_2 images from each year into smaller regions and augmenting them through rotation, flipping, and striding.

The first layer applies a set of 16 convolutional filters, each with a kernel size of 3×3, extracting spatial features while preserving the input dimensions through "same" padding. Following this convolutional operation, the spatial dimensions of the feature maps are reduced using a 2×2 max pooling layer. This pooling step helps the model capture lower-resolution representations of the input while reducing computational complexity. To mitigate overfitting, a dropout layer is applied after pooling, randomly deactivating 25% of the neurons during training.

Subsequently, a second convolutional layer is introduced, applying 32 filters and using the same kernel size of 3×3 with ReLU activation and "same" padding. The output from this layer undergoes an upsampling operation, restoring the spatial dimensions to their original size. Finally, the output layer produces a single-channel result

Table 1: Performance comparison of various CNN configurations. Bold values indicate the best-performing model configuration. The values in parentheses correspond to those obtained over 1000 different training-validation-test splits.

Patch Size	Network Depth	Filters	Inputs	R^2	R^2_0	% p-values < 0.05
32	2	8	IWI	0.20	-0.04	30%
			IWI, wind	0.29	0.03	40%
		16	IWI	0.37	0.24	50%
			IWI, wind	0.34	0.10	40%
	3	16	IWI	0.51	-0.10	50%
			IWI, wind	0.50	-0.10	60%
		32	**IWI**	**0.57 (0.56)**	**-0.06 (-0.01)**	**80% (69.7%)**
			IWI, wind	0.53	-0.15	70%
16	2	8	IWI	-0.14	-0.17	10%
			IWI, wind	-0.13	-0.23	20%
		16	IWI	-0.11	0.03	10%
			IWI, wind	-0.10	-0.02	20%
	3	16	IWI	0.18	0.03	30%
			IWI, wind	0.12	-0.26	20%
		32	IWI	0.14	-0.20	30%
			IWI, wind	0.14	-0.10	30%

Table 2: Performance comparison of various alternative models configurations.

Model	R^2	R^2_0	% p-values < 0.05
forest	-0.28	-0.47	100%
linear	0.09	0.00	100%
VGG16	0.24	0.47	50%

using a 3×3 convolutional filter with ReLU activation, ensuring the predicted patch matches the input dimensions.

The model was trained using the Adam optimizer and a custom loss function designed to handle missing values through a masked mean squared error approach. Patches of size 32×32 were processed in batches of 1,000, and the training was conducted over 10 epochs. Test data comprised 20% of the total available patches, randomly split from the dataset. To prevent leakage, the test dataset was geographically isolated, ensuring that no patches from the same region were included in both the training and test sets, even when derived from images taken in different years.

This architecture was selected after testing multiple configurations, including CNNs of various depths and with varying number of convolutional filters, VGG16-based models, random forest regressors and other CNN models incorporating additional input variables such as wind data from ERA5 - the global climate reanalysis dataset produced by the European Centre for Medium-Range Weather Forecasts (ECMWF) - as summarized in Tables 1 and 2. We considered incorporating wind as an additional input because wind influences the dispersion and transport of NO_2, affecting observed column densities independently of emission rates. However, adding wind did not enhance predictive performance, suggesting that the NO_2-IWI relationship is robust even without accounting for wind variability. Alternative CNN configurations utilizing either smaller or larger patch sizes than those presented in Table 1 demonstrated suboptimal performance. Models with smaller patch sizes often failed to converge, while those with larger patches had R_0^2 values comparable to R^2, indicating a lack of meaningful predictive power.

The tables demonstrate that the best-performing model is a CNN with a patch size of 32 pixels, 3 layers, 32 filters, and IWI as the sole input. This model achieves an R^2 of 0.57, an R_0^2 of -0.06, and 80% of p -values below 0.05. For each model, R^2 values were computed over 10 different training-validation-test splits of the data. The value presented in the table is the mean of these R^2 scores across the 10 splits. R_0^2 represents the average R^2 obtained by training models on shuffled NO_2 data (i.e., randomizing the target variable). This process destroys meaningful relationships, and values close to 0 are expected, indicating poor performance. Specifically, for each training-validation-test split, the NO_2 images were shuffled 10 different times, and a model was trained on each shuffled dataset. The reported R_0^2 value is the mean over all these shuffled cases. Comparing R^2 with R_0^2 highlights how much better the model performs relative to a random baseline. The percentage of p -values minor than 0.05 represents the proportion of splits where the model's performance was statistically significant compared to the benchmark R_0^2. A p-value below 0.05 indicates that the model's performance is unlikely to have occurred by chance. A higher percentage here signifies more consistent and reliable model performance across the different data splits.

To ensure the robustness of our results, the model with the best-performing configuration was trained a greater number of times, that is on 1000 different training-validation-test splits. For each split, the NO_2 images were shuffled 100 times to calculate the benchmark R_0^2. Across these splits, the model achieved a R^2 of 0.56, a R_0^2 of -0.01 and a p-value below 0.05 in 69.7% of cases compared to the random-shuffling benchmark (values in parenthesis in Table 1).

Scenario modeling and forecasting

To project NO_2 column densities in 2030, a year widely used as a benchmark for emission reduction commitments in international agreements, we employed a linear interpolation of IWI data over time for each pixel. Given that the dataset spans from 1990 to 2024, this interpolation was used to estimate IWI values for 2030, which were then input into our trained model to predict future NO_2 column densities. This method was chosen for its simplicity, providing an estimate of NO_2 column densities under a scenario where socioeconomic trends continue along their historical trajectory.

For country-level column densities in 2030, we aggregated predicted NO_2 column densities across all pixels corresponding to populated areas, i.e., those where IWI data were available. The change in column densities between 2019 (as from TROPOMI observations) and 2030 (model predictions) was then computed for each country. To quantify uncertainties in the column densitie predictions, we propagated errors from two main sources: uncertainties in the linear regression parameters used to project IWI forward in time and the measurement uncertainty of TROPOMI NO_2 observations in 2019.

These changes in NO_2 column densities between 2019 and 2030 were directly compared to the change in Nationally Determined Contributions (NDCs) for total greenhouse gas (GHG) emissions, reported in CO_2 equivalents (CO_2e).

This approach is based on the assumption that NO_2 column densities are strongly correlated with emissions of other GHGs, particularly CO_2, due to their shared emission sources in human activities such as industrial processes, transportation, and energy production. Because NO_2 is co-emitted with CO_2 in combustion and industrial activities, it serves as a useful proxy for tracking broader pollutant emissions. Prior research has demonstrated the viability of using NO_2 as an estimator of CO_2 emissions [27].

However, this approach has inherent limitations. Emission ratios between NO_2 and other pollutants can vary across sectors and regions, influenced by fuel type, combustion efficiency, and pollution control technologies. While NO_2 provides a valuable spatial and temporal indicator of emission trends, care must be taken in directly translating NO_2 variations to CO_2e reductions. Further refinements, including sector-specific adjustments and additional environmental factors, could improve the accuracy of this forecasting framework.

Source data

IWI estimates can be downloaded at https://dataverse.harvard.edu/dataset.xhtml?persistentId=doi:10.7910/DVN/9DINV4 [28]; TROPOMI L2 data can be downloaded upon registration at https://disc.gsfc.nasa.gov/datasets/S5P_L2__NO2____HiR_2/summary?keywords=TROPOMI [29]

References

[1] Geiges, A., Nauels, A., Parra, P.Y., Andrijevic, M., Hare, W., Pfleiderer, P., Schaeffer, M., Schleussner, C.-F.: Incremental improvements of 2030 targets insufficient to achieve the paris agreement goals. Earth System Dynamics **11**(3), 697–708 (2020) https://doi.org/10.5194/esd-11-697-2020

[2] Roelfsema, M., Soest, H.L., Harmsen, M., Vuuren, D.P., Bertram, C., Elzen, M., Höhne, N., Iacobuta, G., Krey, V., Kriegler, E., Luderer, G., Riahi, K., Ueckerdt, F., Després, J., Drouet, L., Emmerling, J., Frank, S., Fricko, O., Gidden, M., Humpenöder, F., Huppmann, D., Fujimori, S., Fragkiadakis, K., Gi, K., Keramidas, K., Köberle, A.C., Aleluia Reis, L., Rochedo, P., Schaeffer, R., Oshiro, K., Vrontisi, Z., Chen, W., Iyer, G.C., Edmonds, J., Kannavou, M., Jiang, K., Mathur, R., Safonov, G., Vishwanathan, S.S.: Taking stock of national climate policies to evaluate implementation of the paris agreement. Nature Communications **11**(1), 2096 (2020) https://doi.org/10.1038/s41467-020-15414-6

[3] Huggins, A.: In: Mayer, B., Zahar, A. (eds.) The Paris Agreement's Article 15 Mechanism: An Incomplete Compliance Strategy, pp. 99–110. Cambridge University Press, Cambridge, UK (2021)

[4] Jongaramrungruang, S., Matheou, G., Thorpe, A.K., Zeng, Z.-C., Frankenberg, C.: Remote sensing of methane plumes: instrument tradeoff analysis for detecting and quantifying local sources at global scale. Atmospheric Measurement Techniques **14**(12), 7999–8017 (2021) https://doi.org/10.5194/amt-14-7999-2021

[5] Archer, D., Brovkin, V.: The millennial atmospheric lifetime of anthropogenic co_2. Climatic Change **90**(3), 283–297 (2008) https://doi.org/10.1007/s10584-008-9413-1

[6] Lan, X., Thoning, K.W., Dlugokencky, E.J.: Trends in globally-averaged CH_4, N_2O, and SF_6 determined from NOAA Global Monitoring Laboratory measurements (2025). https://doi.org/10.15138/P8XG-AA10 . https://doi.org/10.15138/P8XG-AA10

[7] Lange, K., Richter, A., Burrows, J.P.: Variability of nitrogen oxide emission fluxes and lifetimes estimated from sentinel-5p tropomi observations. Atmospheric Chemistry and Physics **22**(4), 2745–2767 (2022) https://doi.org/10.5194/acp-22-2745-2022

[8] Munir, S., Mayfield, M., Coca, D.: Understanding spatial variability of no2 in urban areas using spatial modelling and data fusion approaches. Atmosphere **12**(2) (2021) https://doi.org/10.3390/atmos12020179

[9] Scheibenreif, L., Mommert, M., Borth, D.: Estimation of Air Pollution with Remote Sensing Data: Revealing Greenhouse Gas Emissions from Space (2021). https://arxiv.org/abs/2108.13902

[10] Zhang, Z., Wang, J., Lu, W.: Exposure to nitrogen dioxide and chronic obstructive pulmonary disease (copd) in adults: a systematic review and meta-analysis. Environmental Science and Pollution Research **25**(15), 15133–15145 (2018) https://doi.org/10.1007/s11356-018-1629-7

[11] Liu, F., Beirle, S., Joiner, J., Choi, S., Tao, Z., Knowland, K.E., Smith, S.J., Tong, D.Q., Ma, S., Fasnacht, Z.T., Wagner, T.: High-resolution mapping of nitrogen oxide emissions in large us cities from tropomi retrievals of tropospheric nitrogen dioxide columns. Atmospheric Chemistry and Physics **24**(6), 3717–3728 (2024) https://doi.org/10.5194/acp-24-3717-2024

[12] Smits, J., Steendijk, R.: The international wealth index (iwi). Social Indicators Research **122**(1), 65–85 (2015) https://doi.org/10.1007/s11205-014-0683-x

[13] Pettersson, M., Kakooei, M., Ortheden, J., Johansson, F., Daoud, A.: Time series of satellite imagery improve deep learning estimates of neighborhood-level poverty in africa, pp. 6165–6173 (2023). https://doi.org/10.24963/ijcai.2023/684

[14] Redi, T.: Systematic review of air pollution in ethiopia: Focusing on indoor and outdoor sources, health, environment, economy impacts and regulatory frameworks. F1000Research **13**, 1385 (2024) https://doi.org/10.12688/f1000research.157621.1 . [version 1; peer review: awaiting peer review]

[15] Adeniyi, O., Folarin, O.: Industrialization, finance, and urbanization in africa. Journal of African Development **26**(1), 137–164 (2025) https://doi.org/10.

5325/jafrideve.26.1.0137 https://scholarlypublishingcollective.org/psup/african-development/article-pdf/26/1/137/2181083/jafrideve_26_1_137.pdf

[16] Air Pollution – East Africa (ASAP-East Africa), A.S.A.: A systems approach to air pollution – east africa: Synthesis report. Technical report, UK Department for International Development (2020). Accessed: 2025-02-07. https://assets.publishing.service.gov.uk/media/5eb16ef186650c4356562f91/ASAP_-_Air_Pollution_in_East_Africa_-_Synthesis_Report.pdf

[17] Qu, H., Suphachalasai, S., Thube, S., Walker, S.: South africa carbon pricing and climate mitigation policy. Selected Issues Paper SIP/2023/040, International Monetary Fund (2023). Accessed: 2025-02-07. https://www.imf.org/en/Publications/selected-issues-papers/Issues/2023/06/26/South-Africa-Carbon-Pricing-and-Climate-Mitigation-Policy-South-Africa-535220

[18] South Africa, G.: South Africa's First Nationally Determined Contribution under the Paris Agreement: Updated September 2021. Accessed: 2025-02-07 (2021). https://unfccc.int/sites/default/files/NDC/2022-06/South%20Africa%20updated%20first%20NDC%20September%202021.pdf

[19] Ellerbeck, S.: These start-ups are helping to make life in the sahel more sustainable. World Economic Forum (2023). Accessed: 2025-02-07

[20] Bichler, R., Bittner, M.: Comparison between economic growth and satellite-based measurements of no2 pollution over northern italy. Atmospheric Environment **272**, 118948 (2022) https://doi.org/10.1016/j.atmosenv.2022.118948

[21] European Space Agency: Sentinel-5P Nitrogen Dioxide Level 2 Product Readme File. Issue 2.4 (2023). https://sentinel.esa.int/documents/247904/3541451/Sentinel-5P-Nitrogen-Dioxide-Level-2-Product-Readme-File

[22] Geffen, J., Eskes, H., Compernolle, S., Pinardi, G., Verhoelst, T., Lambert, J.-C., Sneep, M., Linden, M., Ludewig, A., Boersma, K.F., Veefkind, J.P.: Sentinel-5p tropomi no_2 retrieval: impact of version v2.2 improvements and comparisons with omi and ground-based data. Atmospheric Measurement Techniques **15**(7), 2037–2060 (2022) https://doi.org/10.5194/amt-15-2037-2022

[23] Judd, L.M., Al-Saadi, J.A., Szykman, J.J., Valin, L.C., Janz, S.J., Kowalewski, M.G., Eskes, H.J., Veefkind, J.P., Cede, A., Mueller, M., Gebetsberger, M., Swap, R., Pierce, R.B., Nowlan, C.R., Abad, G.G., Nehrir, A., Williams, D.: Evaluating sentinel-5p tropomi tropospheric no_2 column densities with airborne and pandora spectrometers near new york city and long island sound. Atmospheric Measurement Techniques **13**(11), 6113–6140 (2020) https://doi.org/10.5194/amt-13-6113-2020

[24] Geddes, J.A., Martin, R.V., Bucsela, E.J., McLinden, C.A., Cunningham, D.J.M.: Stratosphere–troposphere separation of nitrogen dioxide columns from the tempo geostationary satellite instrument. Atmospheric Measurement Techniques **11**(11), 6271–6287 (2018) https://doi.org/10.5194/amt-11-6271-2018

[25] United Nations Framework Convention on Climate Change: Nationally Determined Contributions (NDCs). https://unfccc.int/process-and-meetings/the-paris-agreement/nationally-determined-contributions-ndcs. Accessed: March 31, 2025

[26] United Nations Framework Convention on Climate Change: Nationally Determined Contributions (NDCs) Registry. https://unfccc.int/NDCREG. Accessed: March 31, 2025

[27] Fuentes Andrade, B., Buchwitz, M., Reuter, M., Bovensmann, H., Richter, A., Boesch, H., Burrows, J.P.: A method for estimating localized co_2 emissions from co-located satellite xco_2 and no_2 images. Atmospheric Measurement Techniques **17**(3), 1145–1173 (2024) https://doi.org/10.5194/amt-17-1145-2024

[28] Pettersson, M.B.: Temporal neighborhood-level material wealth maps of Africa (1990-2019). Harvard Dataverse. Accessed: March 31, 2025 (2023). https://doi.org/10.7910/DVN/9DINV4 . https://doi.org/10.7910/DVN/9DINV4

[29] Copernicus Sentinel-5P (processed by ESA): TROPOMI Level 2 Nitrogen Dioxide total column products, Version 02. European Space Agency. Accessed: March 31, 2025 (2021). https://doi.org/10.5270/S5P-9bnp8q8 . https://doi.org/10.5270/S5P-9bnp8q8