Empowering Environmental Governance with Satellite Data: Piloting a Global Randomized Control Trial *

December 31, 2025

Abstract

Limited access to reliable air quality information in many high-polluting low- and middle-income countries constrains citizens' ability to hold regulators accountable for pollution. This paper introduces a scalable, low-cost system that delivers daily satellite-based $PM_{2.5}$ alerts to influential X (Twitter) users across five candidate cities. Our approach identifies central users through network analysis, classifies them with large language models, and randomly assigns targeted notifications via city-specific X accounts. This design enables us to measure how pollution information diffuses through local online networks and how different types of actors respond to environmental data. Preliminary evidence from the pilot cities shows meaningful engagement with the X alerts, suggesting that timely pollution information can stimulate public attention. Within one month of pilot launch, the alerts reached approximately 70,000 impressions across the pilot cities, with the highest reach in Kolkata and Kanpur. Our findings provide early insight into how real-time environmental information can activate civic pressure and potentially influence regulatory incentives. Future work will assess whether sustained exposure to these alerts can lead to measurable improvements in air quality and environmental governance.

Keywords: air pollution, information dissemination, network analysis

*UChicago Energy and Environment Lab

1 Introduction

Air pollution causes seven million premature deaths annually, making it a significant but often invisible threat (, WHO). To build the regulatory capacity needed to tackle pollution, reliable public data must reach engaged and motivated stakeholders. However, at least one billion people live in countries where their national government does not monitor air pollution, and even more live in countries where pollution data is not made public (OpenAQ, 2025).

This data gap is most severe in high-polluting low- and middle-income countries: 80% of life years lost to $PM_{2.5}$ occur in Asia (EPIC, 2023). Yet, only 7% of Asian governments publicly release this information. These same countries often lack the financial resources to install and maintain monitoring infrastructure.

While prior research has demonstrated the importance of monitoring technology, its broader adoption remains limited by financial constraints and political economy barriers. Jha and Nauze (2022) find that when U.S. embassies began monitoring local air pollution, host countries experienced substantial reductions in $PM_{2.5}$ levels, lowering the risk of premature death for over 300 million people. However, the U.S. State Department has recently announced its plan to dismantle these monitors—discontinuing access to reliable, reference-grade data in nearly three dozen countries, thereby exacerbating the data gap. Considering this retreat, scalable, low-cost alternatives for pollution monitoring and public dissemination are more crucial than ever.

Historically, expensive ground monitors have been the primary source of air quality data. But recent advances in technology enable real-time, spatially granular measurement of local pollution, even in regions with historically sparse monitoring. We propose leveraging this novel monitoring technology to implement a global, targeted information dissemination cam-

paign to provide reliable and timely pollution data to local stakeholders. This innovative intervention could potentially be rolled out as a randomized controlled trial to measure its causal impact on pollution awareness, regulatory action, and long-term air pollution levels. Our intervention aims to empower the public to engage more effectively with the governance process and take meaningful steps to mitigate air pollution and climate change. In addition, making headway in addressing the pollution crisis could serve as an effective gateway for broader environmental action.

2 Literature Review

We hypothesize that the impacts of public and easily accessible environmental data, and a notification system reliant on the same, will differ greatly depending on (i) existing data availability and access; (ii) underlying institutional and governance structures; and (iii) income levels. Importantly, a global scope would enable us to transcend the geographical boundaries of traditional experimental investigations and test whether results from existing studies scale up across regions with different cultures, institutions, and governance regimes. This approach significantly enhances the external validity of our findings, strengthening their practical applicability and generalizability. Furthermore, it allows us to directly address environmental justice concerns by measuring differences in impact between high- and low-income countries. This is particularly crucial as low-income nations often face severe air pollution yet remain understudied in environmental research.

Previous studies have shown that making pollution data easily accessible to the public can significantly improve regulation of environmental violations Buntaine et al. (2024); Konar and Cohen (1997). However, existing studies are limited in terms of their geographical scope, tending to focus on either the United States or China. This limited geographical scope means there is little to no evidence on how governance structures and information provision to the

public can interact to foster meaningful change in environmental regulation. An exception to this is work by Jha and Nauze (2022), who found that a United States embassy air quality monitoring and information program that broadcasts real-time air quality readings across 40 cities led to reductions of fine particulate matter ($PM_{2.5}$) in host cities. Our work builds on this research by expanding analysis beyond cities with U.S. consulates and by examining the mechanisms through which public information can reduce environmental damage.

Furthermore, studies have primarily focused on reducing information asymmetry for the regulator with the explicit aim of improving enforcement targeting. While this is an important facet for strengthening environmental monitoring, it discounts the role of civil society in fostering meaningful change from the ground up. In this context, we hope to build on Buntaine et al. (2024), who conducted an RCT in China to study how citizen appeals to regulators about a firm's environmental violations affect regulatory behavior. Comparing private citizen appeals to public social media complaints, they found that complaints on social media reduced firms' violations by 60 percent and triggered a 12 percent decline in sulfur dioxide emission concentrations during the study period. The paper conducts a careful study of the mechanisms and concludes that public appeals were effective in mitigating environmental damages because they shifted political incentives for regulators away from economic growth and towards avoiding public unrest.

We combine elements from Buntaine et al. (2024) and Jha and Nauze (2022) to analyze the global impact of the roles of different civil society actors and how their effects differ across various government systems. Echoing Jha and Nauze (2022), we have a global scope; echoing Buntaine et al. (2024), our treatments target regulatory and civil society bodies in both private and public settings to identify the mechanisms through which pollution notifications prompt regulatory action.

More broadly, this work contributes to the literature on the role of monitoring in regulatory behavior. Finan et al. (2017) have reviewed this literature in developing country settings and list three channels through which monitoring can improve governance: (i) improved information to regulatory bodies, (ii) deterrence for bad actors, and (iii) empowering citizens to demand and obtain better services. They note that the effectiveness of past interventions is quite heterogeneous, and usually hinges on whether those receiving the information have the incentives and ability to act on new information. They further highlight the increasing role of technology in promoting transparency and accountability. Our eventual RCT will aim to test the importance of each of the three channels in pollution monitoring. At the same time, our comparative politics lens is consistent with Finan et al. (2017)'s takeaway that effectiveness of monitoring interventions is context-specific, depending on institutions and state capacity, and information alone may not be enough in environments with insufficient stakeholder buy-in.

3 Data and Methods

To test whether targeted information can influence stakeholder engagement, we designed and deployed a pilot notification system in five cities: Chiang Mai, Guatemala City, Kanpur, Kigali, and Kolkata. Implementing our pilot study required coordinating several logistical components. We first required a $PM_{2.5}$ dataset to determine when notifications should be triggered. Second, we needed to identify appropriate stakeholders to receive our notifications. Third, we developed a delivery system to disseminate notifications. Finally, we established a tracking system to scrape X and capture engagement outcomes.

3.1 Satellite-Derived Global $PM_{2.5}$ Forecasts

We use global forecast estimates of fine particulate matter ($PM_{2.5}$) from the Copernicus Atmosphere Monitoring Service (CAMS), delivered operationally through our partner Plume-

Labs. CAMS produces a globally consistent atmospheric composition product by assimilating a wide range of satellite observations (including aerosol optical depth and trace-gas retrievals), meteorological data, and emissions inventories (anthropogenic, biogenic, dust, volcanic, and fire emissions) into the ECMWF Integrated Forecasting System (IFS). The IFS is extended with aerosol-chemistry modules that simulate emission, transport, secondary aerosol formation, deposition, and aerosol microphysical processes, enabling a physically grounded representation of $PM_{2.5}$ mass concentration at the surface.

CAMS provides global surface-level $PM_{2.5}$ fields on a regular $0.4^\circ \times 0.4^\circ$ grid (roughly 40 km resolution), with hourly values and 5-day forecasts updated twice daily (00 and 12 UTC). Although these gridded estimates are available at an hourly temporal resolution, PlumeLabs aggregates the hourly values into daily mean $PM_{2.5}$ concentrations, which we then utilize to generate same-day forecast alerts for participating cities in our global notification system. Daily averaging reduces short-term model noise, smooths meteorological variability, and aligns the data with public-health–relevant exposure indicators commonly used in epidemiological and regulatory contexts.

This global, forecast-capable $PM_{2.5}$ product enables consistent comparisons across countries and supports early-warning notifications in regions with sparse or nonexistent ground monitoring networks. However, as with other global aerosol models, CAMS estimates have limitations: the 40 km spatial resolution tends to smooth urban hot spots, and model-derived $PM_{2.5}$ can exhibit systematic biases (often overestimation) relative to in-situ measurements, particularly during extreme pollution events or in regions with complex aerosol processes.

3.2 Finding Stakeholders: X Users Extraction

Our goal was to send location-specific notifications to users who were likely to engage with, and potentially act on, information about air pollution. This required both identifying local

users in each pilot city and classifying the types of stakeholders they represented.

To identify the X users residing in each of our five pilot cities, we used snowball sampling, starting with an initial set of stakeholders and then identifying their followers and retweeters to find more relevant users. We repeated this process until we could not find any additional users who met our criteria. To be considered "eligible", each user had to satisfy the following conditions:

1. Have at least 100 followers
2. Have the desired city as their location in their profile

We used the first condition to filter out users with little to no influence on X, and the second to ensure the extracted users lived in the desired city. For example, if a user posts "Kolkata is a great city" but has "India" as a location (or no location at all), then we do not code these users as residing in Kolkata. We used regular expressions to capture all possible variations of the city's name, including its aliases.

To assemble an initial pool of users to feed into the snowball sampling, we searched for posts that mention the city's name (or one of its aliases) over several months. The extracted records contain the tweet's metadata, such as creation date, and the author's information (username, provided location, etc). We then retained only users who met the above criteria (more than 100 followers and the city as their profile location).

To continue with the snowball sampling, for any given user i, we followed this procedure:

1. Extract up to 1,000 followers: Due to computational constraints, we extract up to 1,000 followers per user. We then retain the eligible ones (with at least 100 followers of their own and residing in the city) and repeat the entire process for each. That is, we get each follower's tweets, retweeters, and followers.

2. Extract up to 1,000 tweets: We filter the tweets only to keep the user's original posts, so retweets are removed. These original tweets are stored and, if reposted by another user, analyzed further.

3. Extract, for each tweet j, up to 250 retweeters: For each post, we extract all up to 250 retweeters. Again, if the retweeters meet the necessary conditions, we repeat the snowballing process for each of them (obtain their tweets, followers, and retweeters).

As the name suggests, the number of users to extract grows exponentially until no more eligible users are found. However, the time required to complete the snowballing process depends entirely on X's API rate limits and our computational capacity to parse and store the data.

Table 1: Extracted users per city

City	**Extracted users**
Chiang Mai	4,203
Guatemala City	4,380
Kanpur	5,241
Kigali	9,599
Kolkata	21,024

The following section describes how we identify the most influential users after identifying all eligible people in each city.

3.2.1 Finding Central Users in Each City's Network

To make the intervention more effective, we targeted the most influential users in each city. For this component of the project, we relied heavily on Network Theory. We can think of each user as a node (or vertex) in a network and map the relationships between them (these would be the edges). For example, if $user_A$ follows $user_B$, we code a directed edge from $user_A$ to $user_B$. This can also be represented as $user_A \xrightarrow{follows} user_B$. Another type of relationship

would exist if $user_A$ also retweeted $user_B$, which would be described as $user_A \xrightarrow{retweeted} user_B$. Therefore, we can define each city as a directed graph $G_{city} = (V, E)$.

During the snowball sampling, we map each city's network in Amazon Web Serrvices' (AWS) Neptune by creating a node for each user and edges between them (either when they follow or retweet someone). The "follows" edge is straightforward, but the "retweeted" edge has an additional consideration: a user can retweet another user several times across different posts. For example, $user_A$ may have retweeted several of $user_B$'s tweets. We incorporate this feature by assigning a weight to each "retweeted" edge, where the weight is the number of times a given user retweeted another user. Such weights play an important role in identifying each city's most important users.

In Figure 1, we show a sub-network for each city, selecting a sample of 2,500 nodes and their "retweets" edges. As expected, each city's subnetwork is different.

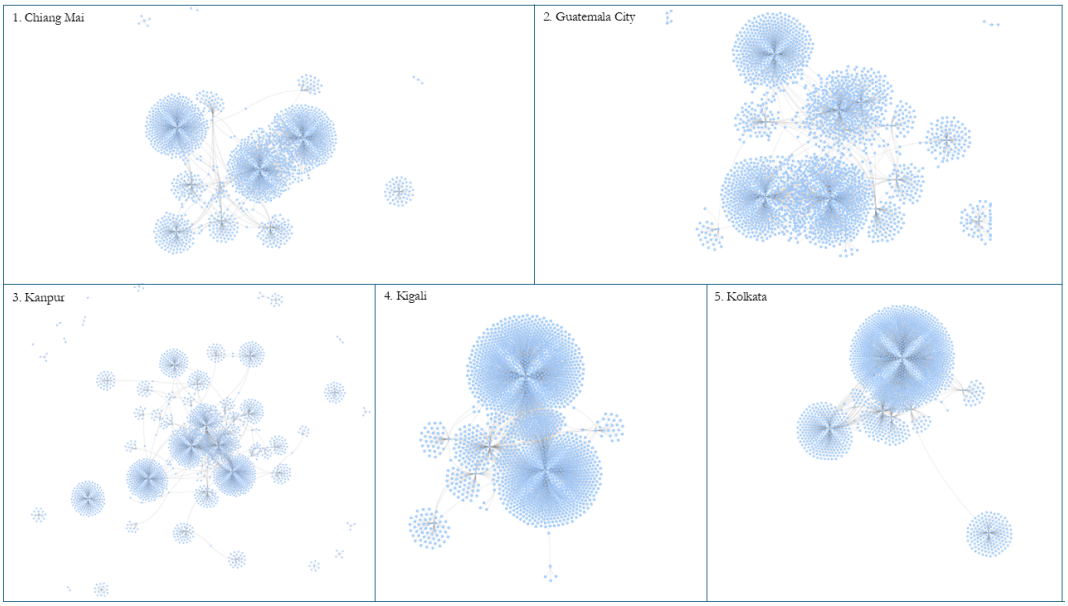

Figure 1: Sub-networks for Pilot Cities with 2,500 nodes - "Retweeted" Edges

We then apply the Weighted PageRank centrality algorithm to each city's subgraph (defined as $G_{city} = (V, E)$) to identify the most influential users in each city. Note that the Weighted Page Rank algorithm is a variation of the PageRank Algorithm, which Page et al. (1999) developed at Google to rank websites in their search engine results. The core intuition is that a website's importance is determined by two factors: the number of incoming links (or edges) and the quality of those links. Therefore, the PageRank of page u can be defined as:

$$PR(u) = (1 - d) + d \sum_{v \in B_u} \frac{PR(v)}{L(v)}$$

Where $B(u)$ is the set of pages that point to u, $PR(u)$ and $PR(v)$ are the rank scores of page u and v respectively, and d is a dampening factor usually set to 0.85. Note that u denotes a given node in the network (in our case, a node represents a given user).

Xing and Ghorbani (2004) extend this concept by assigning higher ranks to more popular nodes in Weighted PageRank (Xing and Ghorbani, 2004). In normal PageRank, page u distributes its "influence" equally among its outlink pages, whereas in Weighted PageRank, the nodes with higher rank will capture more of page u's influence. Note that to page u, an "inlink" is a URL of another webpage that contains a link pointing to it (an incoming edge). Similarly, an "outlink" is a link appearing in u that points to another webpage (an outgoing edge). Weighted PageRank then introduces two weights which are applied in the algorithm:

$W^{in}_{(v,u)}$ is the assigned weight of $link(v, u)$, calculated based on the number of incoming edges to u and the number of incoming edges of all reference pages of page v.

$$W^{in}_{(v,u)} = \frac{I_u}{\left(\sum_{p \in R(v)} I_p\right)}$$

where I_u and I_p represent the number of incoming edges of both u and p, and $R(v)$ represents the list of all "reference pages" of v.

$W^{out}_{(v,u)}$ is the assigned weight of $link(v,u)$, calculated based on the number of outgoing edges of u and the number of outgoing edges of all reference pages of page v.

$$W^{out}_{(v,u)} = \frac{O_u}{\left(\sum_{p \in R(v)} O_p\right)}$$

where O_u and O_p represent the number of outgoing edges of both u and p, and $R(v)$ represents the list of all "reference pages" of v.

Then, the original Weighted PageRank algorithm is defined as follows:

$$PR(u) = (1-d) + d \sum_{v \in B_u} \frac{PR(v)}{L(v)} W^{in}_{(v,u)} W^{out}_{(v,u)}$$

However, in our context, we only use $W^{out}_{(v,u)}$, which is represented by the number of times $user_v$ retweeted $user_u$. Therefore, we adapt the algorithm as follows:

$$PR(u) = (1-d) + d \sum_{v \in B_u} \frac{PR(v)}{L(v)} W^{in}_{(v,u)}$$

We use both the weighted and unweighted versions of the PageRank algorithm to each city's subgraph, $G_{city} = (V, E)$ in AWS Neptune, depending on the edge type considered for the analysis. When using the "follows" edge, we use the unweighted version of the algorithm, while we use the adapted Weighted PageRank when using the "retweets" edges.

We obtain a score for each user indicating their "influence" on the network. Note that a user that is considered influential in the "retweets" network may not necessarily be considered influential in the "follows" network and vice versa. We then retain the most influential users in each city and classify them using Large Language Models (LLMs), which are covered in detail in the next section.

3.2.2 Classifying Each City's Central Users

Beyond identifying how influential each user was, we also sought to understand what type of user we were engaging with. To do this, we use an LLM to classify each influential user into different categories. We instructed an LLM to analyze the user's description and tweets, which we extracted during the snowballing process, to:

1. Determine if the user is an organization or an individual.
2. Determine whether the user belongs to one of nine categories, depending on whether it is an organization or an individual.
3. Provide an environmental relevance score from one to three, depending on how vocal the user is of environmental issues.

Specifically, we created a prompt that included the user's description and their most recent *original* tweets to enable the model to perform a two-stage classification task. In the prompt, we explicitly defined each category, provided examples, and determined how to solve certain edge cases. We also requested that the model provide a "confidence score" from 0 to 1 for its decision, along with an explanation.

The following list contains a summarized portion of the prompt that explains each category, assuming the user was first classified as an individual.

1. Media Journalism: Journalists, reporters, editors, or independent media creators focused on news or factual content, such as global or local geopolitics or events, breaking news, alerts, etc. It should not include lifestyle or entertainment content propagated by public influencers.
2. Government: Individuals who are part of or directly affiliated with a government, including non-elected public officials such as diplomats, regulators, policy advisors, and

civil servants, as well as elected or appointed politicians such as presidents, ministers, MPs, mayors, and political party leaders.

3. Activism Advocacy: Individuals primarily engaged in activism, advocacy, or campaigning for causes, aiming to influence society or government actions without being part of its administration. This includes activists, NGO leaders or staff, grassroots organizers, and individuals focused on social, climate, or human rights movements. Also include individuals affiliated with campaign organizations, ideological advocacy groups, or religious accounts that promote faith-based values as a form of advocacy.

4. Academic Research: Researchers, PhDs, postdocs, professors, think-tank staff, or those individuals contributing to the research.

5. Healthcare: Doctors, nurses, medical officers, and public health experts.

6. Tech Entrepreneurship: Founders, engineers, or startup leads, especially in AI/ML, energy, or climate tech.

7. Education outreach: Educators, science communicators, teachers, or public-facing STEM professionals.

8. Public Influencer: Celebrities, athletes, creators, or social media figures with broad public followings. These are built around a persona and focus on entertainment or lifestyle content, rather than factual news dissemination

9. Other: None of the above apply, or not enough information to classify.

Essentially, the LLM is solving a binary classification problem to decide whether the user is an "individual" or an "organization", and a multi-class classification problem. To evaluate the model's performance, we conducted two rounds of evaluation, computing accuracy metrics for both classification problems and adjusting the prompt as needed.

For the first validation round, we randomly selected 200 users, 40 from each city, and had two annotators perform the classifications by hand. Each analyst could review the users' biographies and tweets. Also, the annotators were not allowed to check the profile on X or view any profile pictures, as the LLM was unable to do so. If the two annotators did not agree on a classification, an additional classification round with two additional annotators was conducted until consensus was reached. At the end of the annotation process, we compiled a human-annotated ground-truth dataset to evaluate the model.

We then calculated the Weighted F1 Score for both classification tasks using two LLMs, *gpt-4.1-mini* from OpenAI (referred to as GPT) and *gemini-2.5-flash* from Google (referred to as Gemini). The scores are displayed in Table 2, where both models perform similarly on binary classification, and GPT outperforms Gemini in category classification.

Table 2: F1 Weighted Scores - First Validation Round

Task	**gpt-4.1-mini**	**gemini-2.5-flash**
Binary classification	0.96	0.96
Category classification	0.68	0.48

Given GPT's superior performance, we chose it as the default model for classification. The first validation round allowed us to identify areas of improvement in our prompt based on the model's mistakes. After improving our prompt, we conducted a second validation round focused exclusively on classifying central users.

In the second validation round, for each pilot city, we randomly selected 20 users from the pool of influential users. This yielded a representative sample of 100 users, with the annotation process performed exactly as in the first round. This process resulted in a second ground-truth dataset that we used for evaluation. After further refinements, we noticed improvements in both classification tasks, particularly in the category classification. The

Weighted F1 Scores are provided in Table 3 below.

Table 3: F1 Weighted Scores - Second Validation Round

Task	**gpt-4.1-mini**
Binary classification	0.97
Category classification	0.80

Figure 2 shows the distribution of classification categories across the pilot cities (where we selected the top 100 influential users per city).

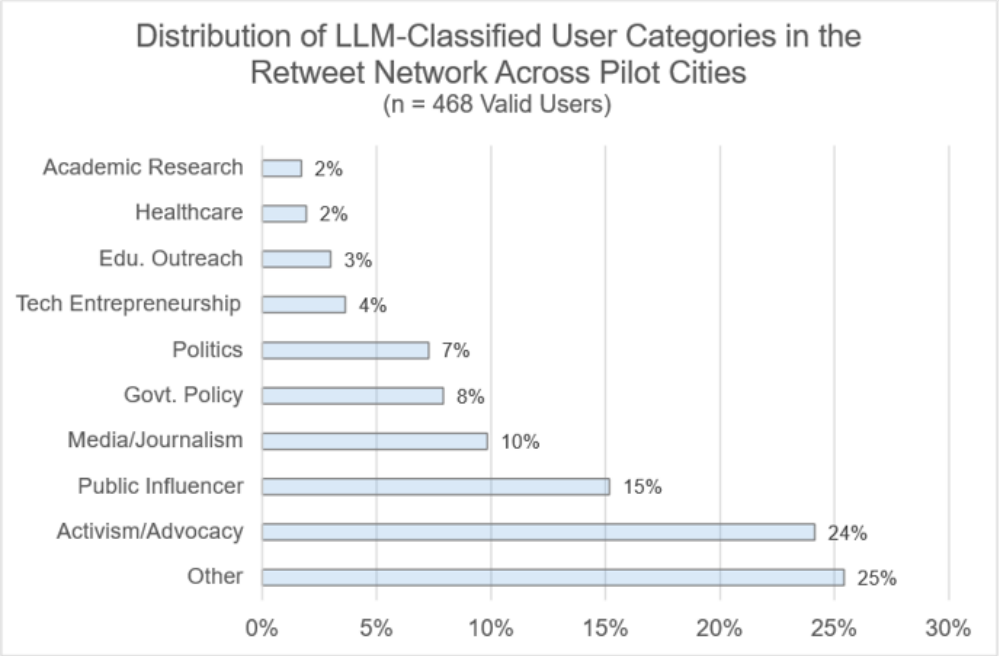

Figure 2: Distribution of Classification Categories Across Pilot Cities

The next section will cover the pollution alerts system we built on X, along with the logic of when to tag the central users within each city.

3.3 Pollution Alerts and Information Dissemination

To disseminate information to the stakeholders we identified, we created a notification delivery system on X. Specifically, we established one X account per city to post frequent updates about local air pollution conditions. We also sought to understand what types of content might be most effective. Therefore, for each post, we randomly assigned the following characteristics at the time of posting:

1. Tweet type: Determines if the tweet will be a "general pollution update" or a "comparison" tweet.
2. Media: Determines if a plot will be included or not (the plots will also change depending on the tweet type).
3. Language: Determines if the tweet will be in English or in the city's local language.

Figure 3 displays the logic for posting non-health-related tweets. Note that we also tag a random subset of central users from each city when certain conditions are fulfilled.

The different tweet types, along with the tagging logic, are discussed below.

Pollution Tweets

In a "pollution tweet", we initially provide the daily $PM_{2.5}$ level and indicate if the WHO standard has been surpassed. Also note that if we choose to attach media to the post, we randomly select either a line plot of daily $PM_{2.5}$ levels or a bar plot comparing today's pollution to the WHO standard.

We check for a "spike" in pollution by checking if the current $PM_{2.5}$ levels meet the following criteria. If the following conditions are fulfilled, we say there is a pollution spike and we tag a random subset of influential users from the city:

1. Higher than the city's $PM_{2.5}$ National Standards

2. Higher than the previous 15 days of $PM_{2.5}$ (e.g., 17th October, higher than every point in the 15 days)

3. Should be at least 10% higher than the maximum from $t - 15$ to $t - 5$.

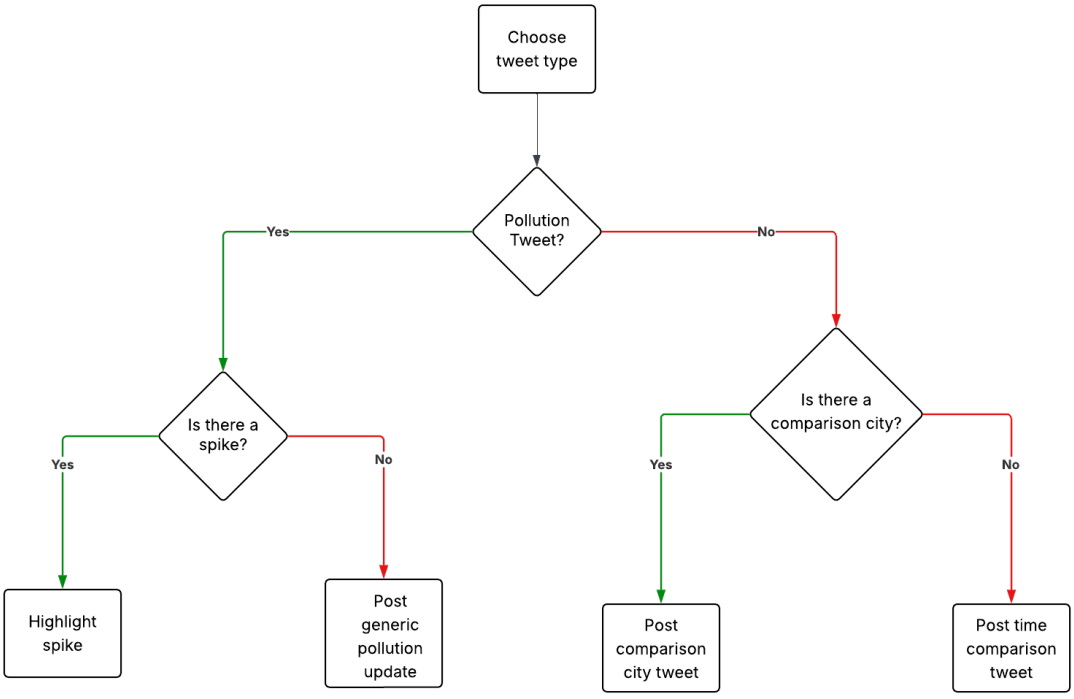

Figure 3: Posting Logic for Non-Health-Related Tweets

Figure 4 below shows a spike in pollution in Kanpur, and the plot focuses on the last couple of weeks. If there is no pollution spike, we do not focus on the last 15 days or simply compare the current day's $PM_{2.5}$ levels to the WHO Standard.

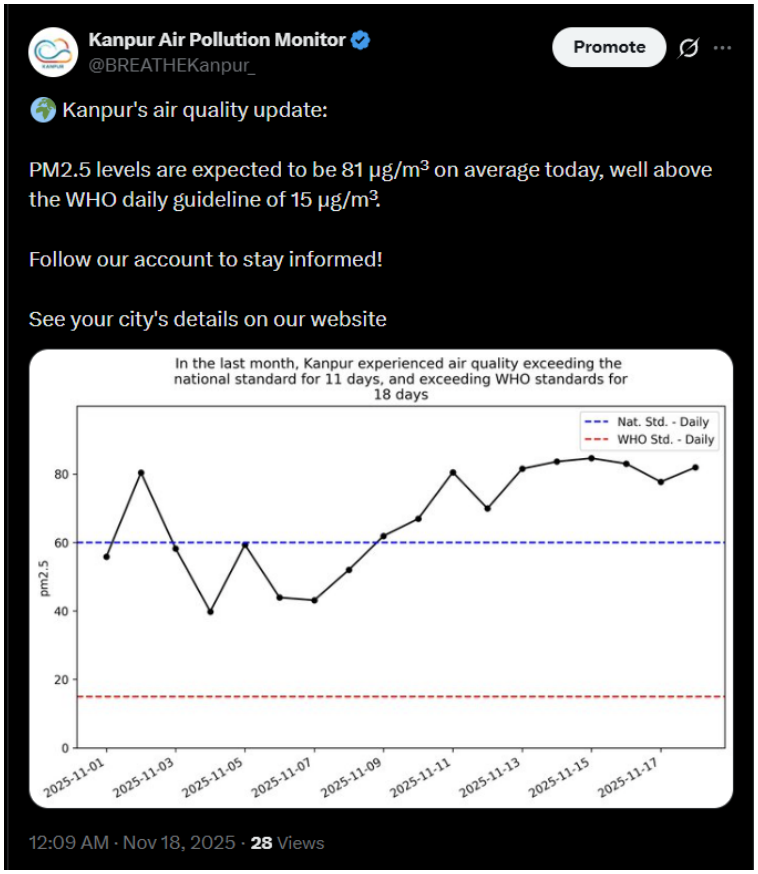

Figure 4: Spike Pollution Tweet for Kanpur

Comparison Tweets

There are two choices for comparison tweets:

1. City comparison: We compare our target city with the cleanest of the four comparison cities we picked. We post the city comparison tweet and tag users if the target city's $PM_{2.5}$ is at least 10% higher than that of the cleanest comparison city. Otherwise, we switch to a time-based comparison tweet.

2. Time comparison: In this setting, we compare today's $PM_{2.5}$ levels with pollution averages over different time windows (monthly, quarterly, six-month, annual, and "around the same time last year"). We again pick the lowest of these averages to make the comparison and tag users if the city's $PM_{2.5}$ is at least 10% higher than that of the cleanest period. If this is not fulfilled, we do not tag users, but we still post the tweet

regardless.

Figure 5 below shows a city comparison tweet for Chiang Mai, where Phuket was the cleanest comparison city at the time.

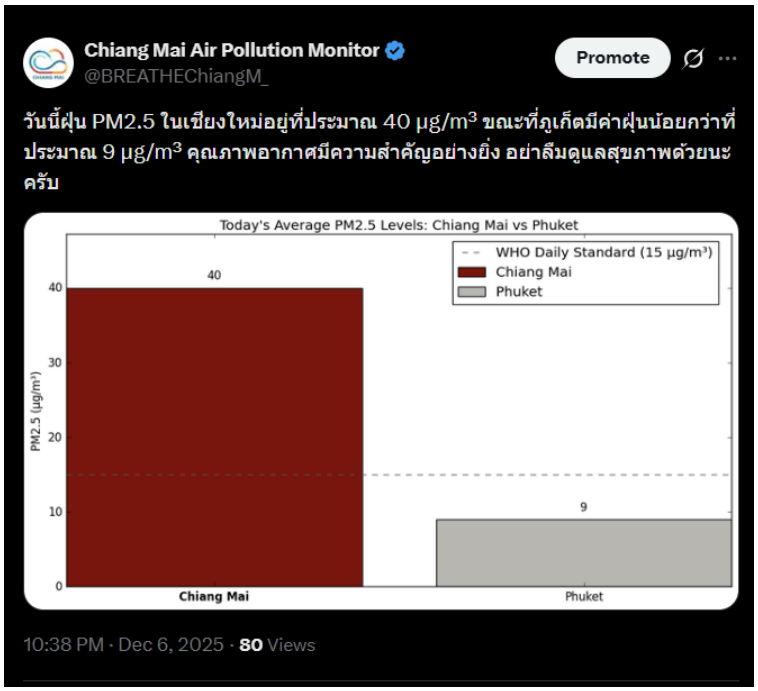

Figure 5: City Comparison Tweet for Chiang Mai

To test a health-based framing, we replaced the usual pollution-level post on Sundays with a tweet reporting the equivalent number of cigarettes smoked due to air pollution in each city. We also randomly decide whether a plot should be attached to the post and whether the tweet's text will be translated. In this scenario, influential users will always be tagged.

Figure 6 shows an example for Kolkata, with the graph showing the total number of cigarettes smoked during the last thirty days.

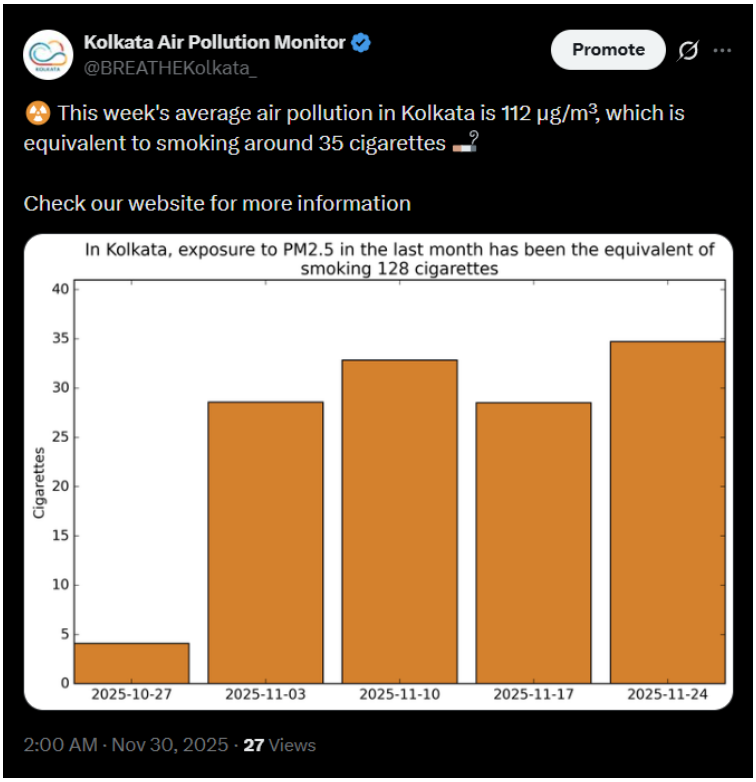

Figure 6: Health Tweet for Kolkata

3.3.1 Website

To accompany our global notification system, we developed a public-facing website, which provides accessible visualizations and contextual information on air quality and its health impacts. The site presents daily and annual $PM_{2.5}$ levels for participating cities using clear charts and interactive graphics, enabling users to track pollution trends over time. In addition to air-quality data, the platform includes evidence-based summaries of the health risks associated with $PM_{2.5}$ exposure, links to relevant scientific research, and guidance for reducing pollution-related harm. The website is designed to serve as both a communication tool and an educational resource, making complex environmental and health information understandable to the general public, policymakers, and community organizations. By offering pollution insights alongside concise explanations of their implications, the platform reinforces the project's broader goal of empowering environmental governance through ac-

cessible, data-driven information.

3.4 Tracking Engagement

We created a system from scratch to track several engagement metrics for the alerts we send. The first component of the system consists of recording relevant information from the alert every time it is sent. For example, we store the tweet's language, corresponding city, author id, and whether media is included, if users are tagged, among others.

The second component focuses on tracking and recording engagement metrics periodically after a twelve-hour delay. After the initial delay, we run the engagement tracking process every 3 days for the next 2 weeks. We track and update engagement metrics in Dynamo DB, a NoSQL Database system provided by AWS. The engagement metrics include likes, views, retweets, replies, and quotes.

This engagement system is fundamental since it allows us to perform analyses on how the alerts perform, especially to see which type of content is more attractive and tends to spark more discussion.

4 Results

So far, we have sent 425 pollution alerts over the last three months across all cities in our pilot study, disseminating pollution information to the general population. Table 4 shows the total engagement metrics per city when considering two boosted posts for Kolkata and Kanpur, while 5 shows the metrics without them.

Table 4: Aggregate Engagement Metrics by City with boosted tweets - Count

City	# tweets	Views	Likes	Retweets
Kigali	91	2,191	21	6
Guatemala	85	2,431	16	7
Chiang Mai	83	4,895	18	10
Kanpur	84	23,332	45	7
Kolkata	82	36,753	52	14

Table 5: Aggregate Engagement Metrics by City without boosted tweets - Count

City	# tweets	Views	Likes	Retweets
Kigali	91	2,191	21	6
Guatemala	85	2,431	16	7
Chiang Mai	83	4,895	18	10
Kanpur	83	3,415	23	7
Kolkata	81	3,607	37	13

The number of tweets per city is similar, but not exactly the same since we did not post tweets daily in the early stages of the project. Moreover, Kanpur and Kolkata have significantly more views because we used X's "boosting" feature for two specific posts. The boosted posts account for the large number of views and engagement metrics for both cities. As expected, the average number of views and likes for Kanpur and Kolkata are higher due to the engagement boosting. This is shown in Table 6 below. Table 7 shows the same average engagement metrics but removing Kolkata's and Kanpur's boosted posts.

Table 6: Average Engagement Metrics by City (with Boosted Tweets) - Mean

City	# tweets	Views	Likes	Retweets	Bookmarks	Replies
Kigali	91	24.08	0.23	0.07	0.02	0.00
Guatemala	85	28.60	0.19	0.08	0.04	0.00
Chiang Mai	83	58.98	0.22	0.12	0.01	0.00
Kanpur	84	277.76	0.54	0.08	0.06	0.00
Kolkata	82	448.21	0.63	0.17	0.04	0.01

Table 7: Average Engagement Metrics by City (without Boosted Tweets) - Mean

City	# tweets	Views	Likes	Retweets	Bookmarks	Replies
Kigali	91	24.08	0.23	0.07	0.02	0.00
Guatemala	85	28.60	0.19	0.08	0.04	0.00
Chiang Mai	83	58.98	0.22	0.12	0.01	0.00
Kanpur	83	40.18	0.27	0.16	0.06	0.00
Kolkata	81	43.99	0.45	0.17	0.04	0.01

So far, even though our alerts are being seen by the population on Twitter and there seems to be some baseline engagement through likes and reposts, there is almost no discussion triggered by such alerts (if we considered replies as a proxy of discussion). The following figures show how the engagement view has evolved over time. Figure 7 shows how Chiang Mai has a relative upward trend in average weekly views, but Kigali and Guatemala City have stagnated.

Figure 8 shows both Kolkata and Kanpur's weekly average views, with a clear spike in engagement when the boosting feature was activated for both cities. The effect lingered for a couple of days, but not more than one week. Figure 9 shows the average weekly views for both cities when removing the "boosted" tweets, and reveal a similar trend when comparing to Guatemala City and Kigali. However, the average engagement for both Indian cities is higher than Guatemala City and Kigali, but lower than Chiang Mai.

One of the main challenges of the project has been increasing our account's reach and in-

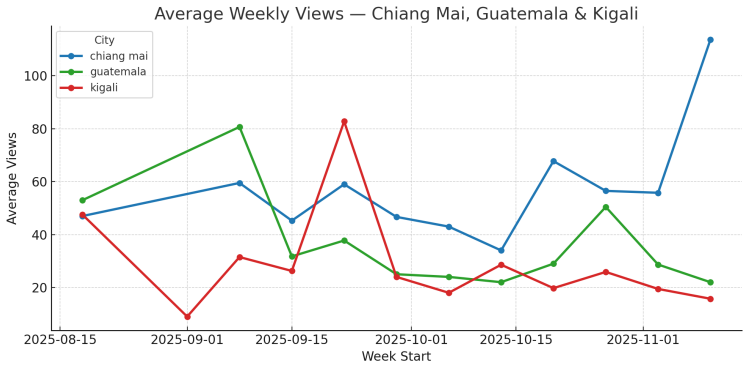

Figure 7: Average Weekly Views - Chiang Mai, Guatemala City, and Kigali

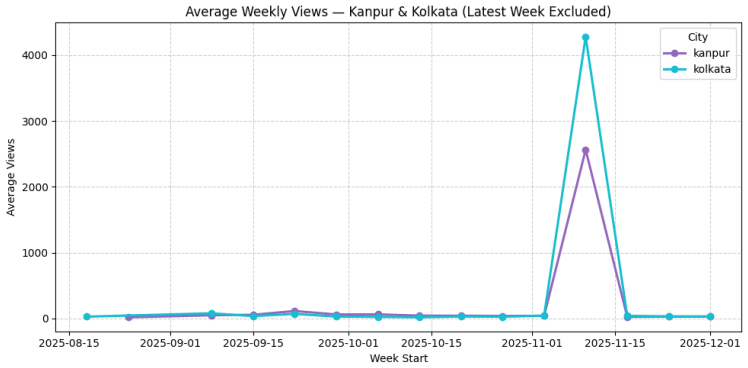

Figure 8: Average Weekly Views - Kolkata and Kanpur with Boosted Tweets

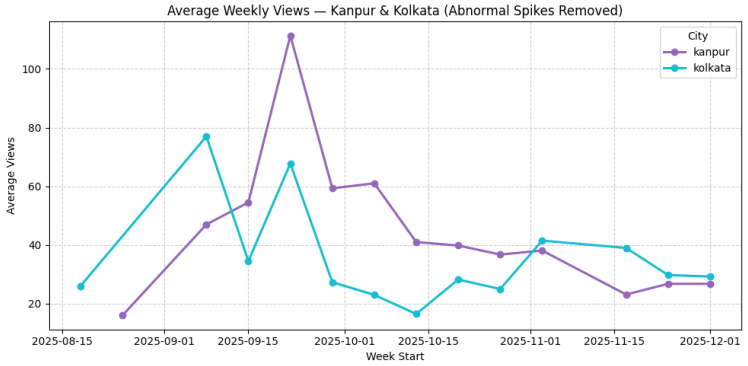

Figure 9: Average Weekly Views - Kolkata and Kanpur without Boosted Tweets

fluence organically. X's algorithm heavily penalizes "repetitive content and posts" and, even despite our best efforts to diversify our content, we seem to fall short on engagement and reach due to X's algorithm. One of the solutions revolves around the use of X's "boosting"

feature to quickly gain engagement and increase the account's reach and reputation. However, scaling the experiment to a larger number of cities using the same approach becomes much more complex.

5 Conclusions and Future Research

This study demonstrates the feasibility of building a global, low-cost, and scalable system to deliver timely pollution information to influential local actors via social media. By combining satellite-assimilated $PM_{2.5}$ forecasts with city-level network mapping, central user classification, and targeted alerts, we create a platform capable of reaching stakeholders who shape public conversations about air quality.

We sent more than 420 pollution alerts across the five cities on X, providing useful information about air quality in such regions. Using diversified content and sporadically using X's "boosting" feature, we managed to reach around 70,000 impressions in total, with notably more engagement in Kolkata and Kanpur.

Looking ahead, an important avenue for expanding this work is to shift from simply disseminating information toward activating civic and regulatory responses. As our accounts grow in reach and credibility within each city, we plan to test strategies for mobilizing followers to discuss pollution spikes, amplify public concern, and apply pressure on local regulators to improve air-quality management. Future research will evaluate the effectiveness of these engagement-driven interventions, study how institutional and political contexts condition their impact, and assess whether sustained public attention can translate into lasting improvements in air quality and environmental governance worldwide.

References

Buntaine, Mark T, Michael Greenstone, Guojun He, Mengdi Liu, Shaoda Wang,

and Bing Zhang, "Does the squeaky wheel get more grease? The direct and indirect effects of citizen participation on environmental governance in China," *American Economic Review*, 2024, *114* (3), 815–850.

EPIC, "Air Quality Life Index (AQLI): Annual update 2023," April 2023.

Finan, Frederico, Benjamin A Olken, and Rohini Pande, "The personnel economics of the developing state," *Handbook of economic field experiments*, 2017, *2*, 467–514.

Jha, Akshaya and Andrea La Nauze, "US Embassy air-quality tweets led to global health benefits," *Proceedings of the National Academy of Sciences*, 2022, *119* (44), e2201092119.

Konar, Shameek and Mark A Cohen, "Information as regulation: The effect of community right to know laws on toxic emissions," *Journal of Environmental Economics and Management*, 1997, *32* (1), 109–124.

OpenAQ, "Open Air Quality Data: The Global Landscape," 2025.

Page, Lawrence, Sergey Brin, Rajeev Motwani, and Terry Winograd, "The PageRank Citation Ranking: Bringing Order to the Web.," Technical Report 1999-66, Stanford InfoLab November 1999. Previous number = SIDL-WP-1999-0120.

(WHO), World Health Organization, "Air pollution," 2025.

Xing, W. and A. Ghorbani, "Weighted PageRank Algorithm," in "Proceedings. Second Annual Conference on Communication Networks and Services Research, 2004." 2004, pp. 305–314.