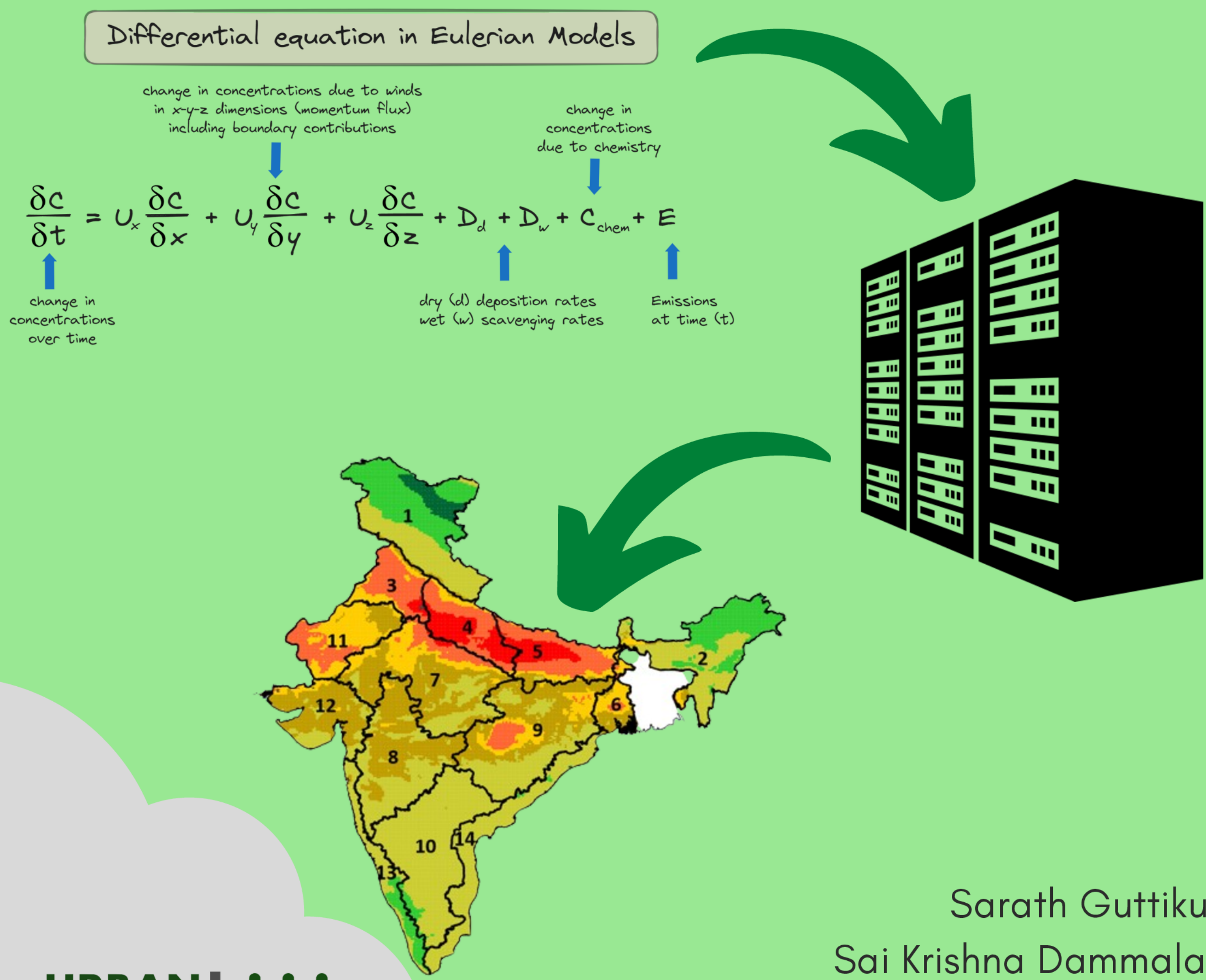
LANDSCAPE REVIEW OF AIR QUALITY MODELING IN INDIA



Sarath Guttikunda
Sai Krishna Dammalapati
Nishadh Kalladath

SIM-air working paper series #50-2024

URBAN emissions .info

UrbanEmissions (UEinfo) was founded in 2007 with the vision to be a repository of information, research, and analysis related to air pollution. UEinfo has four objectives: (1) sharing knowledge on air pollution (2) providing science based air quality analysis (3) promoting advocacy and raising awareness on air quality management and (4) building partnerships among local, national, and international air-heads.

All our publications are accessible @ www.urbanemissions.info/publications

We would like thank Dr Pallavi Pant (Health Effects Institute, Boston, USA) for helping us with the initial phase of journal articles tally and review.

Send your questions and comments to simair@urbanemissions.info

Short Story

“All the models are wrong, only some are useful”
– George Box, Mathematician

A lot of progress is made in modeling India’s air quality from emissions to pollution to health impacts to scenario analysis, at regional and urban scales. In this working paper, we reviewed the progress made by the modeling community, identified some research gaps, and proposed some line items to extend India’s air quality modeling efforts (content-wise and institutionally). The landscape review is limited to the level of research activities and did not investigate the depth of the research activities.

Applications of classical chemical transport models are limited in India, primarily due to operational knowledge gaps. Classical models, with their detailed representations of atmospheric processes and emissions, offer the comprehensive understanding necessary to tackle our complex air pollution issues. Therefore, ensuring that researchers have the resources, training, and data needed to engage with this modeling should remain a priority.

While reduced complexity models can effectively support policy dialogues by providing rapid assessments and insights, the core research questions related to air quality and its impacts can only be fully addressed when the barriers to classical modeling are removed.

More networking opportunities can bridge the gap between local and global researchers, fostering collaboration that enhances the quality and relevance of research.

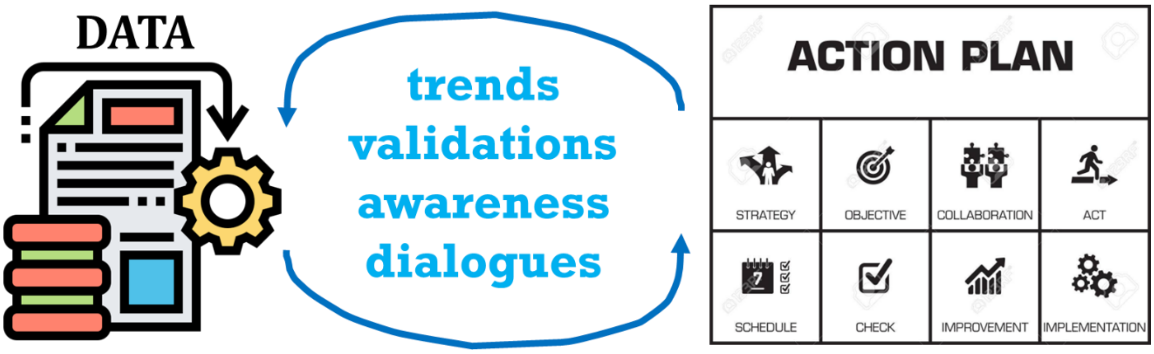
A model-intercomparison exercise is due for India and the Indian Subcontinent, to officially launch a representative emissions inventory and an air pollution modeling framework. By strengthening these foundational capabilities, the scientific community can produce more robust analysis that inform and assess air quality management strategies.

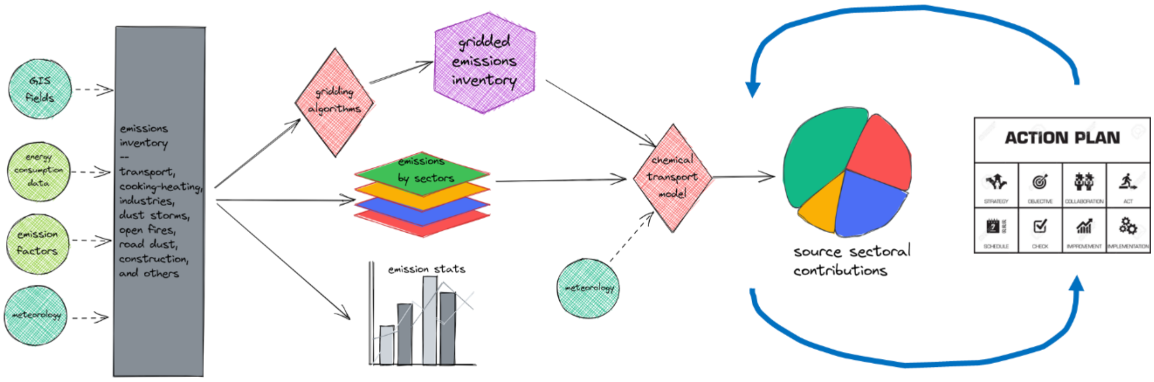
Better engagement through communication platforms with the public and public bodies can ensure that communities understand the research methods and their implications, fostering a more informed dialogue and driving meaningful, evidence-based policy decisions.

1. Air Quality Modeling (AQM)

"All the models are wrong, only some are useful"
– George Box, Mathematician

Air quality modeling (AQM) is "data management" and the goal of this exercise is to support clean air action planning and implementation (Fowler et al., 2020). This data provides us with the baseline to understand the trends that inform us if pollution levels are going up or down over time. Various forms of this data can be used to raise awareness among the public and the public bodies for informed activism and decision making. This data from monitoring networks and modeling platforms provides us with the knowledge to validate the progress made or not made from various emissions management programs in space and time. All this data further strengthens the dialogue between data generators, data consumers, policy makers, and the public.





In this working paper, we refer to AQM following the classical modeling path presented in the above figure. AQM is an involved exercise, in need of a lot of data and a lot of computational power from the first step of emissions modeling to evaluating clean air action plans. When all the components of the exercise come together, forms of input and output data can support various stages of air quality management process.

Key components of AQM

1. Emissions modeling: The output of this exercise is a multi-pollutant emissions inventory that includes model-ready speciation for both aerosols and gases. The level of speciation, particularly for volatile organic compounds (VOCs), is determined by the chemical mechanism selected for the chemical transport modeling phase. This inventory must comprehensively represent all anthropogenic activities within the chosen regional or urban airshed, as well as natural sources, organized by spatial (model grids) and temporal (seasonality and diurnal variations) dimensions. For the anthropogenic emissions inventory, there are no user forums or operational training resources available. While fundamental equations and some coarse-level methodologies are openly accessible, much of the work is localized, relying on available input fields.

2. Meteorological modeling: The output from this exercise is a three-dimensional database encompassing all the essential meteorological parameters, including wind speed and direction, temperature, pressure, precipitation, and mixing heights (potential boundary layer heights). Typically, the advanced Weather Research and Forecasting (WRF) model is employed for this purpose, offering multiple physics parameterizations to cater to various geographical contexts around the world. This model can generate meteorological fields at a wide range of spatial resolutions, down to 1 km, and even 100-500 m in super-high resolution downscaling mode. Additionally, there is a well-established user forum for the WRF model, featuring a global community of users and mentors who provide support and guidance.

3. Chemical transport modeling (CTMs): This stage integrates outputs from emissions and meteorological models, allowing for the analysis of air quality from various perspectives. It helps identify the sources contributing to air pollution over both spatial and temporal dimensions. Questions such as which hotspots in the airshed require immediate attention and whether the proposed clean air action plans will effectively reduce current pollution levels, can be addressed using CTM outputs. Furthermore, these results are utilized to assess the health impacts of excessive air pollution and serve as inputs for cost-benefit analysis of avoidable air pollution under "what-if" scenarios.

4. Ambient and Emissions Monitoring: This is a crucial input to AQM exercise, as this provides the necessary information to validate CTMs and subsequently validate all the inputs from the emissions and meteorological models. Any discrepancies in the emissions and meteorological data will show deviations between ground measurements and model results. While data from the regulatory monitoring systems is preferred, the emerging technologies in the form of sensors and algorithms to retrieve data from satellite feeds are also useful. The later datasets are considered unofficial, but a very useful resource to the modeling community.

2. Scope of this Review

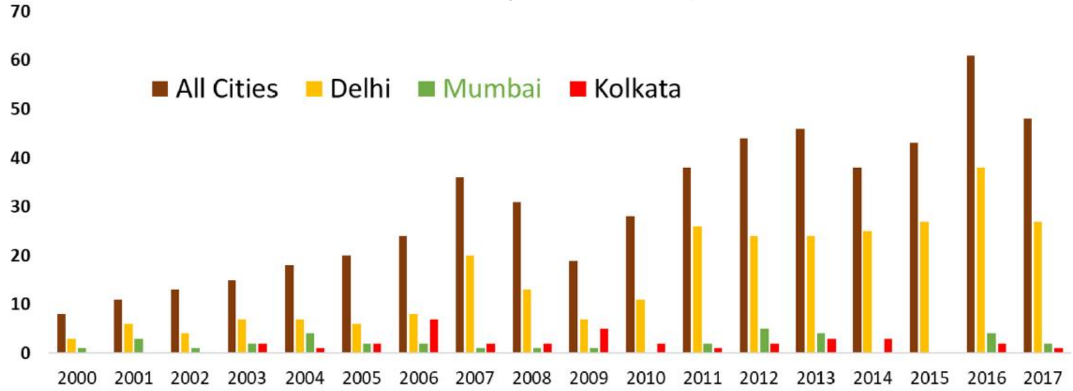
Scope of this working paper is to present the landscape review of research on India's air quality in four core bins – emissions modeling, meteorological modeling, pollution modeling, and monitoring. The health-related epidemiological studies were left out of this assessment, as they form a larger research community (related to medical background) and are covered extensively as part of the global burden of disease assessments. Quality checks were introduced to also exclude studies related to indoor air pollution.

The landscape review is limited to the level of research activities on India's air quality and did not investigate the depth of the research activities.

India is a large and diverse country with a population of approximately 1.4 billion and growing. To provide a sense of urban density and monitoring scale, Delhi's population of 20 million accounts for less than 2% of the total, residing in less than 0.2% of the national land cover, and yet is the most monitored and most studied, and most represented in the media for its air pollution problems (Guttikunda et al., 2023). A large part of the literature ignores the rest of India and as there is little data and studies to quantify the impacts of India's air pollution.

A review presented in (Guttikunda et al., 2014) and (Guttikunda et al., 2019) aimed to understand the nature of air pollution in India by summarizing the number of scientific studies conducted in various Indian cities. The review focused on identifying the sources of air pollution and their contributions to ambient $PM_{2.5}$ and PM_{10} levels. Notably, 70% of the published studies repeatedly focused on just five cities, highlighting a significant concentration of research efforts in a few major urban areas, while many other cities with growing air pollution issues remain understudied.

Number of journal articles published between 2000 and 2017 (from SCOPUS search) with some reference to air pollution research in any Indian city. Figure reproduced from (Guttikunda et al., 2019). Delhi remains the most studied city in India for air pollution.



In this working paper, we extend our analysis to examine all the available research related to air pollution in India published since 2000. This categorization aims to provide a comprehensive overview of the state of air pollution research in India, highlighting the areas that have been well-studied and identifying significant gaps where more research is needed. **Through this review, we aim to understand the broader trends in the analytical techniques commonly used for India's AQM exercises**.

By carefully analyzing these trends, we hope to pinpoint specific research gaps that, when addressed, can guide the development of more focused and impactful studies on air quality. AQM research is crucial not only to advance scientific understanding but also to inform the policymaking process. Ultimately, through this working paper, our goal is to provide guidance to enhance the quality and relevance of air quality studies in India, which can inform effective policy development process to reduce the impact of air pollution on public health and the environment.

3. Data Source

Data collection and analysis was conducted in three stages – first batch in 2014, second in 2018/19 and the third in June 2023. The first batch was presented in (Guttikunda et al., 2014) and the latter two are documented in brief in this working paper.

To quantify the landscape of air quality research in India, we categorize scientific journal articles into those focused on emissions modeling, dispersion modeling, source apportionment and air monitoring (including satellite measurements).

Similar meta-analysis has been conducted in the past, covering various regions and sectors, but they often focus on specific topics, allowing for a deeper understanding of research levels both qualitatively and quantitatively. This process is typically referred to as systematic reviews. For instance, these methods are frequently employed to examine the linkages between air pollution and health outcomes, such as assessing the prevalence of asthma cases in a region based on published research, hospital records, news articles, and other relevant sources. We did not do a systematic review.

Box 1: Keywords used during the initial "SCOPUS" article search

General search was conducted over title, affiliations, abstract, and keywords – information covered under these sub-categories is available for all the articles listed under SCOPUS, for free.

Combination of key words were searched in sets

- INDIA AND {Air Pollution} OR {Air Quality} OR {Source Apportionment}
- INDIA AND {Particulate Matter} OR {Black Carbon} OR {PM2.5} OR PM10 or {sulfur dioxide} or {SO2} or {Nitrogen oxides} or NOx or {Volatile Organic Compounds} or VOC OR {secondary organic aerosols}
- INDIA and CAMx OR CMAQ OR {WRF-Chem} OR PMF OR AERMOD OR ISC3 OR CALPUFF OR HYSPLIT OR FLEXPART OR {GEOS-Chem}
- INDIA and EDGAR OR HTAP OR MICS
- INDIA and {Emissions Inventory}; CITY and emissions
- INDIA and {aerosol optical depth}
- INDIA and {air monitoring} and not indoor
- INDIA and {air } and {receptor modeling}

Combination of the above keywords were varied by replacing the regional name (INDIA) with a city name – India, Agra, Ahmedabad, Allahabad, Amritsar, Aurangabad, Bengaluru, Bangalore, Bhopal, Chandigarh, Chennai, Madras, Coimbatore, Delhi, Dhanbad, Ghaziabad, Mumbai, Bombay, Gwalior, Hyderabad, Indore, Jabalpur, Jaipur, Jamshedpur, Jodhpur, Kanpur, Kochi, Kolkata, Culcutta, Kota, Lucknow, Ludhiana, Madurai, Meerut, Nagpur, Nashik, Patna, Pune, Raipur, Rajkot, Ranchi, Srinagar, Surat, Varanasi, Vijayawada, Vadodara, Visakhapatnam.

Combination of keywords for emissions were varied for specific sources – Household, {vehicle exhaust} or congestion or traffic, resuspended dust road dust or dust storms, power plants or thermal power or power generation, waste burning or garbage burning, open fires or biomass burning, biogenics, brick kilns, coal combustion, cow dung.

Combination of keywords, in the title only was searched for specifc regions - {air } and {indo-gangetic} or {bay of bengal} or {indian ocean} or {arabian sea}

For India's AQM, we limited our search to subjective analysis published in peer-reviewed journals that have a digital object identifier (DOI) listed in the SCOPUS database for the period from 1980 to 2022. While this library does not encompass all published works, it provides a representative selection of high-impact journals, allowing us to effectively capture the research landscape. Although several organizations publish air pollution studies that are often referenced in journal articles, we excluded these reports from our analysis, as they lack the traceability associated with DOI-numbered journal articles. We included both open-access and not-open access (needing subscription) articles.

Several keywords were selected for the initial search to capture a wide range of publications. These keywords were used in various combinations, including subjects of interest, regions such as "India" and city names, specific pollutants, established models for emissions, dispersion, and receptor modeling, as well as emission source regions and types, and keywords capitalized and not. A summary of the keywords employed in this process is presented in Box 1. Additionally, a separate search was conducted using the names of known research organizations to ensure that no relevant studies were excluded.

Box 2: Keywords used for article cateloging

For emissions inventory, we conducted search under two bins – one looking at general emission publications and one looking at sources. We also looked at specific acrnoyms referring to global emission inventories

- Emissions (bin 1) - emissions inventory, emission inventory, emission factors, emission factor, EDGAR, HTAP, GAINS
- Emissions (bin 2) - vehicle emissions, vehicle exhaust, diesel, gasoline, CNG, LPG, power plants, power generation, brick kilns, industrial emissions, coal combustion, cookstoves, road dust, dust storms, open fires, biomass burning, cow dung, biogenics, waste burning, garbage burning – in this case, we looked for minimum 2 hits (under the assumption that for any paper with one source discussed, will surely use the keywords from bin 1)

For air monitoring, we conducted search under two bins, to emphasize the new developments in the use satellite data retreival programs to address the gaps in ground level monitoring

- General monitoring (bin 1) - air monitoring, health impacts, mortality, air quality index, aqi, exposure, aerosol optical depth, aod, satellite
- Satellite monitoring (bin 2) - aerosol optical depth, aod, satellite

For dispersion modeling, we conducted search focusing on the commonly used models, under the assumption that if the paper presents any dispersion modeling results, it is likely to mention the model in the title or the abstract or the keywords

- CAMx, CMAQ, WRF-chem, dispersion modeling, UrBAT, ATMoS, ISC3, AERMOD, CALPUFF, HYSPLIT, GEOS-chem, FLEXPART, meteorological modeling, WRF-CMAQ, WRF-CAMx

For source apportionment, we conducted search for receptor modeling based papers

- PMF, CMB, receptor modeling, source apportionment

The four core areas are the "pillars" of AQM community. Not only the published research on air pollution, but also the practitioners can be broadly binned into these four categories, with groups commonly specializing in ground monitoring, satellite data retrievals, emission measurements, dispersion modeling, chemical analysis for receptor modeling, and data mining, and all of them feeding into the integrated air quality research.

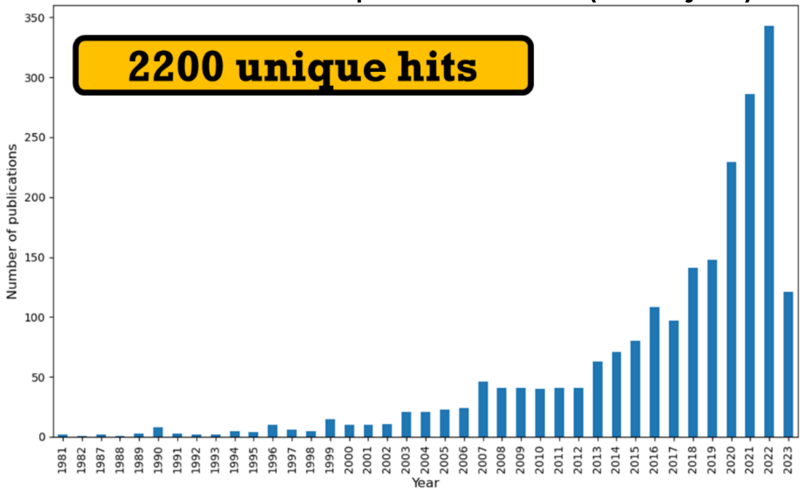
Cataloguing the articles to a specific core area, using keywords was challenging. With some back and forth, several keywords were identified for each of the core areas, listed below, which were further explored during the exercise.

The selection of keywords is based on our understanding of models, data available, data often referred to, and methods employed for analysis under each of these topics. There is a lot of potential to improve these searches in the future. For example, in case of source apportionment, since it includes chemical analysis, any paper which published a chemical analysis is likely to provide a hit, but it is not necessary that the paper is looking to estimate the source contributions following receptor modeling. This review will require deep diving into each of the selected papers at the end of pooling. We limited the analysis to key categories and subject matters only.

4. India's AQM Landscape Review

Between 1980 and 2022, the number of publications related to air pollution research in India has shown a steady increase, reflecting growing interest and attention to this critical issue. This trend is accompanied by an increase in the number of authors per paper, indicating a shift towards more collaborative research efforts over time. On average, the selected papers feature 4.25 authors, with a median of 3.0. These figures highlight the increasingly collaborative nature of research on India's air quality and the importance of interdisciplinary teams in addressing complex environmental challenges.

Number of journal articles published between 1981 and 2023 (from SCOPUS search). More than 60% of these articles were published after 2014 (last ten years).



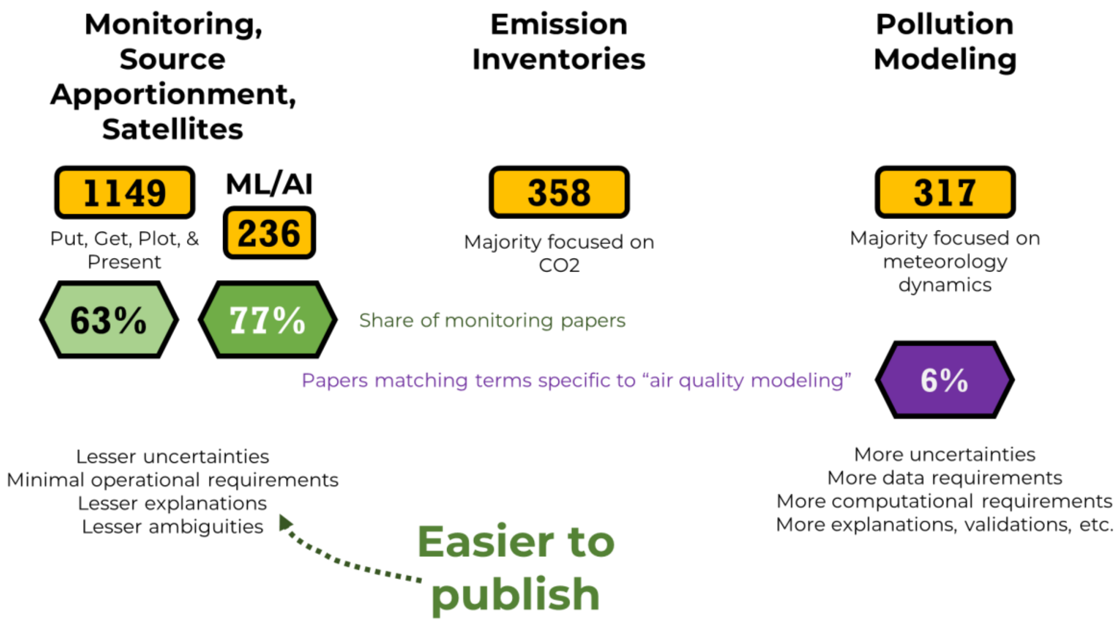
When the papers were categorized by first author nationality, it became evident that studies with a non-Indian first author outpaced those led by Indian authors. While this categorization was based on a subjective assessment of names, which led to the exclusion of 200 papers, the trend remains clear. Papers with non-Indian first authors had an average of 5.6 authors, compared to 3.7 for Indian first-author papers. In 2017, the gap was particularly stark, with non-Indian-led papers averaging 8 authors, while Indian-led papers averaged around 4. **This suggests the need for greater support and networking opportunities for Indian researchers to engage in larger, global research efforts**.

The keywords were analyzed to see growing and changing trends in the air quality modeling community. Here are some of the fields, which showed a growing increase in the occurrences

- Among the key categories, monitoring methods-based papers are the most
- Among the pollutants, PM is the most studied (~40%)
- Among the fuels, diesel is most studied
- Among the meteorological models, WRF is the most utilized (~90%)
- Among the chemical transport models, WRF-Chem is the most utilized

- Among the receptor models, PMF is the most utilized, followed by CMB
- Among the source types – traffic, vehicle exhaust, congestion, power plants, biomass burning, open waste burning, and dust are most mentioned. In general, transport related sources are mentioned the most
- Among the satellites, MODIS and TROPOMI are the most referred and in the recent papers (post 2021) TROPOMI is the most utilized
- Among the monitoring networks, AERONET is the most referred and utilized. Post 2021, low-cost sensor networks are more mentioned
- Among the modeling papers, EDGAR global emissions inventory is the most referred and utilized.

Comparison among the key components of AQM



Many papers published on India's air quality focus on monitoring-based research, accounting for approximately 63% of the pooled total. When papers utilizing newer methods, such as satellite data and information technology (including AI/ML-based approaches), are included, this share rises to 77%. In contrast, only 6% of the papers showed specific discussion on air quality modeling (involving chemical transport models).

This reflects a clear preference for monitoring studies over other aspects of air quality research, such as emissions modeling, dispersion modeling, or receptor-based analysis. A subjective conclusion we draw from this trend is that publishing monitoring-focused journal articles tends to be easier compared to other more complex (and data involved) components of air quality modeling.

Monitoring studies often involves fewer uncertainties and minimal operational requirements, making them simpler to conduct and present. These studies typically follow a straightforward approach: set up monitoring instruments, collect data, plot results, and discuss general trends. This method requires less explanation of methodologies or interpretation of ambiguous results, as it revolves around direct measurements. In contrast, modeling studies require more complex data interpretation, deeper technical expertise, and often involve higher uncertainties due to the need for assumptions and predictions. The relative simplicity and clarity of monitoring-based research, therefore, may contribute to its higher representation in published literature on India's air quality.

This observation regarding the prevalence of monitoring-based studies is in no way intended to downplay the complexity involved in air quality monitoring itself. In fact, monitoring requires a high level of technical expertise in setting up and operating the instruments, as well as rigorous quality assurance and quality control (QA/QC) procedures to ensure the accuracy and reliability of the data. The challenges of maintaining equipment, ensuring proper calibration, and interpreting the data are significant and require substantial knowledge and technical skill. However, the conclusions here are drawn specifically from the path to publication, where monitoring studies tend to face fewer hurdles compared to other more intricate research areas like emissions or chemical transport modeling. **In other words, this also highlights the academic pressures to publish more and faster**.

While the share of pollution modeling papers is significant, a closer look reveals that most of them focus on meteorological modeling using the WRF (Weather Research and Forecasting) model. This can be attributed to several factors: first, the WRF community has made the model highly user-friendly, with easy access to download, compile, and run the software, accompanied by clear instructions and open access to input fields. Despite the high computational demands, the adoption of WRF modeling is relatively straightforward, provided the necessary technical training is available. Additionally, the user community is robust and actively supports new users through forums, helping to resolve technical challenges. The model and its associated methods are well-established and widely recognized, which also makes it an attractive option for researchers seeking a smoother path to publication. These factors collectively contribute to WRF's popularity in air quality research in India.

The adoption rate of the WRF-Chem system, an extension of the WRF model that includes chemical transport modeling, is relatively lower due to the complexity involved in localizing emission inventories and the challenges in explaining uncertainties and deviations in the model's results, particularly in terms of pollutant concentrations and source contributions. These additional layers of complexity make it more difficult for researchers to use and validate the model, leading to its slower uptake compared to the core WRF model. Similar arguments can be applied to adoption of other chemical transport models like CAMx and CMAQ.

While the share of emission inventory papers is significant, a closer look reveals that most of them focus on energy and CO_2 inventories. This is because CO_2 inventories are relatively easier to establish, as the carbon content of the fuel remains constant regardless of the combustion technology used. In contrast, inventories for pollutants such as $PM_{2.5}$, NO_2, CO, and VOCs are far more complex, as their emissions depend on the specific combustion technologies employed and the presence (or absence) of control equipment. These pollutants require detailed knowledge of local industrial practices, fuel types, and technology, making their inventories much harder to compile accurately.

The primary barrier in developing comprehensive emission inventories is the availability of reliable activity data and localized emission factors. While energy data is often more accessible, obtaining detailed information on specific emission sources and their associated technologies presents a challenge. This lack of data creates significant uncertainty in localized emissions estimates, which not only complicates their use in chemical transport models but also makes it challenging to get such studies published. The higher level of ambiguity and the need for extensive validation and explanation of assumptions often result in a more rigorous peer-review process, further slowing the path to publication for these types of inventories.

We are not suggesting that uncertainties should go unacknowledged or that the peer review process should be lenient; rather, we aim to highlight that in LMICs, the road to publication can be particularly challenging. The rigorous scrutiny associated with high-quality research, while essential, often poses additional obstacles for researchers in these regions. Limited access to reliable data and resources can exacerbate the difficulties in addressing uncertainties, making it harder for LMIC researchers to publish their findings and contribute meaningfully to the global discourse on air quality management.

Chemical transport model adoption in India (and China)

The review was extended for generic terms comparing the number of journal articles published on China's air quality. In general, the publication rate is 4-6 times higher than the numbers observed on India's air quality.

	India	China	
Air Quality	**3,475**	**12,123**	
Air Pollution	**5,593**	**22,471**	
Air Quality Modelling	**76**	**315**	
WRF	**601**	**2,318**	
WRF-chem	**139**	**679**	**Full chemistry Eulerian chemical transport models coupled with meteorological models**
CMAQ	**26**	**672**	
CAMx	**7**	**146**	
AERMOD+ISC3+ISCST3	**78**	**38**	
Gaussian	**55**	**72**	**Plume/Gaussian/Lagrangian transport models**
CHIMERE+MOZART+SILAM	**33**	**54**	
GEOS-chem	**69**	**318**	**Global Eulerian chemical transport models, with everything**
GAINS	**22**	**46**	
InMAP	**0**	**2**	**Reduced complexity models for integrated scenario analysis**

This section explored the adoption rates of chemical transport models to study India's and China's air quality. The most used models are listed in the table and the rates are low. The same numbers compared to applications in the US and EU show ratios higher than 30-50.

The global models are mostly institutionalized in the US and EU, and any studies published on India's air quality using those models were also conducted by those groups. While these models are open, for example, the GEOS-chem system on github (https://geoschem.github.io/index.html), is not an easy model to download, compile and run. These systems need a lot of computational power and deep learning, before such models can be housed in India.

Qualitative Comparison of Chemical Transport Models

	Complexity	Ease of Operations	Data Requirement	Computational Requirement	Pollutant Chemistry	Dispersion (Advection)	Level of Details
Box Models	Easy	Easy	Low	Low	Maybe	None	Low
Plume Models	Easy	Medium	Medium	Low	Maybe	Minimum	Low
3D Eulerian Models	High	Complex	High	High	Full	Full	High

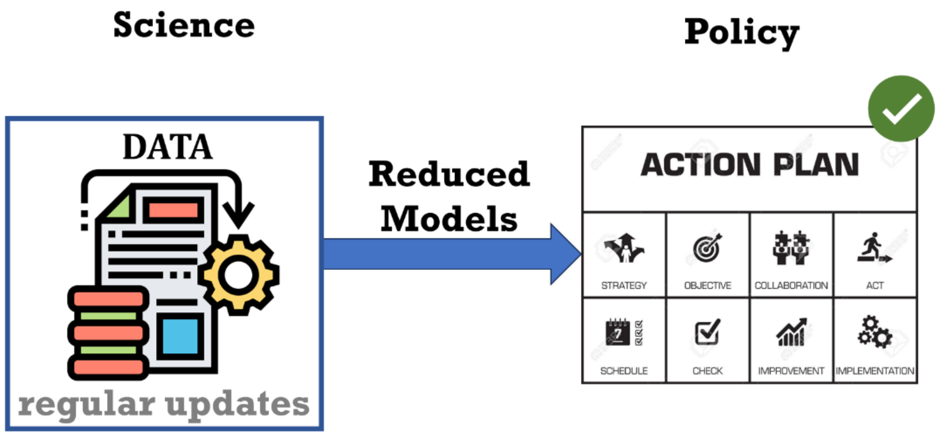
Personnel expertise requirement is a must for all the models

Plume and Gaussian models, such as AERMOD and ISC3, have seen more applications in India than in China, primarily because these models are mandated for environmental impact clearance certification for industries in India. While many of these studies do not pursue the journal publication route, they generate a substantial number of application reports within the country. Typically, an environmental engineer engaged in air quality modeling is more likely to be trained in this suite of models than in the classical regional and global chemical transport models. This focus on plume and Gaussian models reflects the practical requirements of industry compliance, leading to their widespread use in the Indian context.

We are particularly interested in the applications of classical models such as WRF-Chem, CAMx, and CMAQ, as they enable us to explore the mathematical, physical, and chemical aspects of air pollution at various scales. The barriers to the adoption of these models have been discussed in the next section, highlighting the challenges researchers face in utilizing these more complex systems effectively.

Is there a role for reduced complexity models in India?

GAINS and InMAP are integrated models which can go from emissions to concentrations to impacts to scenario analysis via cost-benefit analysis and fall into the category of reduced complexity models. Reduced complexity models serve as a valuable tool in air quality research by taking outputs from classical models and developing a set of matrix functions between emissions and concentrations that allow for rapid assessments of air pollution impacts. These models bypass the role of running the classical models every time an emission scenario is thought of, thus simplifying the modeling process and enabling quicker decision-making, which is particularly beneficial for policymakers and practitioners who need timely information. For example, GAINS (https://iiasa.ac.at/models-tools-data/gains) provides valuable insights into the interplay between air pollution control and greenhouse gas emissions, making it an essential tool for addressing both environmental and public health challenges in the EU. Confidence in using the GAINS system has been built over two to three decades, during which it has received input and enhancements from numerous institutions and experts in the field. This collaborative development process has led to a robust and reliable model that integrates a wealth of data on emissions, costs, and technological options. Similarly, inMAP is making applications in the US and China.



However, it is important to note that the development of these reduced complexity models still relies on initial runs of classical models to generate the necessary data and establish the underlying functions. Additionally, creating accurate emission inventories remains a critical step in this process, as the effectiveness of reduced complexity models hinges on the quality of the input data. Validation is another essential component, ensuring that the simplified models produce reliable and meaningful results. Once these foundational steps are completed, reduced complexity models can be highly effective in providing insights into air quality dynamics.

While reduced complexity models can effectively support policy dialogues by providing rapid assessments and insights, **the core research questions related to air quality, its chemistry and its impacts can only be fully addressed when the barriers to classical modeling are removed**. Classical models, with their detailed representations of atmospheric processes and emissions, offer the comprehensive understanding necessary to tackle complex air pollution issues. Therefore, ensuring that researchers have the resources, training, and data needed to engage with classical modeling should remain a priority. By strengthening these foundational capabilities, the scientific community can produce more robust analysis that inform and assess air quality management strategies.

5. Operational Barriers to AQM

Most of the published literature and the modeling tools for AQM originate from the US and European Union, irrespective of the regions (except for China, in general Northeast Asia covering Japan and South Korea). In the context of low and middle countries (LMICs), it is important to also list the barriers to AQM. The methods and models necessary to support each of the AQM key components are open to the air community. Still, the implementation of these systems to their fullest extent is lagging in LMICs (Gani et al., 2022).

Data limitations (Barrier risk – medium)

This extends across AQM from monitoring to emission inventories. At the start of understanding the pollution problem, we require a lot of monitoring data, just to know where, when, and how much the pollution is, followed by what (source) is contributing to the problem. The monitoring capacity in most of the LMIC's is limited, nascent at best, with a lot of potential to grow. Similar limitations extend to the emission inventories, which is a spatial and temporal representation of the source intensities and a heat map of all the anthropogenic and natural emission activities in the selected regional and urban airshed.

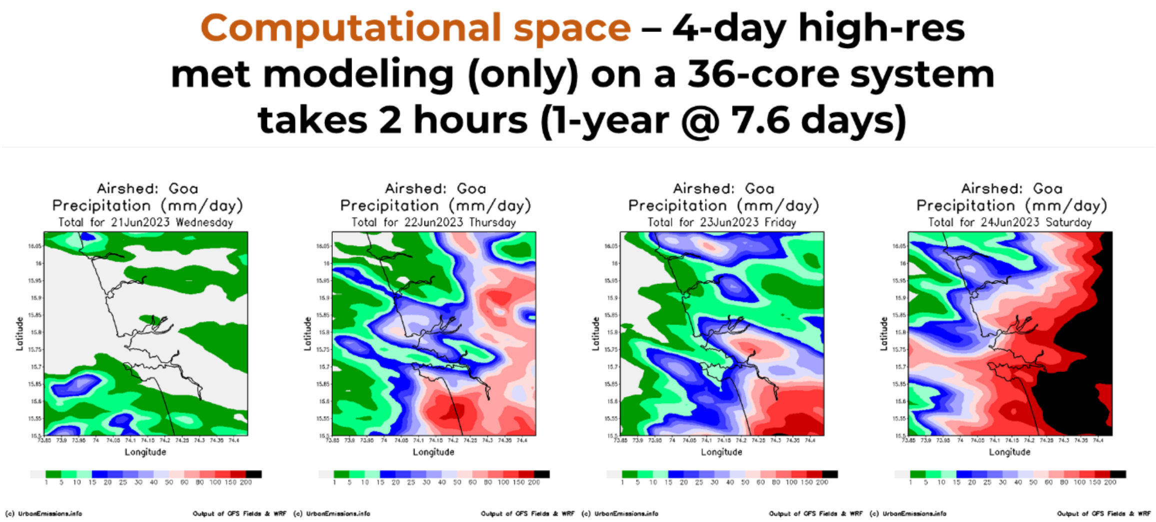
The barrier risk is listed as medium for three reasons

(a) Emerging sensor technology is bridging the gap and increasing the pool of information to study the ground realities (Morawska et al., 2018; Liang et al., 2023)
(b) Emerging science with better algorithms to inverse model satellite retrievals is helping with data for regions where there is no ground monitoring (Holloway et al., 2021) and
(c) Global emissions inventories, nudged and reverse engineered with information from satellite observations, are better at representing the spatial and temporal patterns of various pollutants (as compared what was available 5-10 years ago, with known uncertainty) (Crippa et al., 2023; Crippa et al., 2024; Thunis et al., 2024).

For urban scale assessments, localized emission inventories are preferred to capture the spatial and temporal patterns. Global and regional scale inventories can only provide guidance (and average) assessments over the grids covering the urban areas. The barrier risk for urban-scale emission inventories is high.

Computational space (Barrier risk – low)

Except for the emissions modeling stage, the meteorological and chemical transport modeling stages demand substantial computational power and storage capacity. For instance, the global reanalysis data required to run the WRF model needs at least 1TB of storage for a single year's data, and WRF outputs can range from 100GB to 1TB, depending on the size of the airshed. The figure below illustrates the significant computational requirements involved. These demands are similarly high for chemical transport modeling, varying with the complexity of the models used.



However, the barriers to entry in the current era of information technology are low due to the availability of scalable cloud services with minimal maintenance costs. Shifting away from physical servers, which require significant resources such as housing, electricity, air conditioning, and technical personnel, has significantly reduced the costs of maintaining a computational platform on demand.

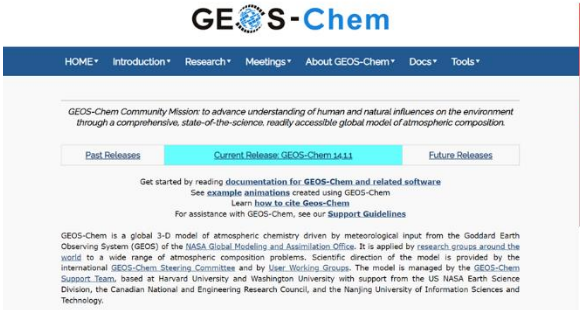
Institutional and training facilities (Barrier risk – high)

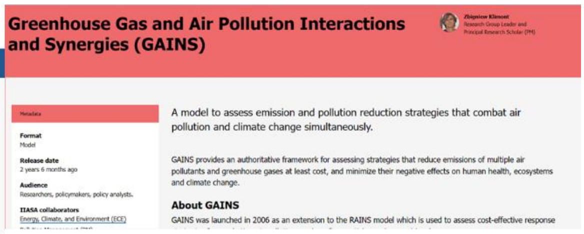
A significant gap exists in institutional housing and training facilities for air quality models and environmental data analysis, especially in LMICs. Many institutions lack the infrastructure and expertise needed to provide hands-on training with the advanced tools (in addition to the computational space) required for information management (from emissions to policy interpretations).

This shortage hampers the development of local capacity, leaving many regions reliant on external expertise or struggling to implement effective air quality solutions. Without proper training facilities, individuals in academia, government agencies, and regulatory bodies find it challenging to develop and apply technical skills needed for effective monitoring, modeling, and policy development.

Moreover, many educational institutions primarily focus on theoretical aspects of environmental science and engineering, as part of the curriculums, often neglecting practical, operational training on models (like WRF meteorological or WRF-chem, CAMx, CMAQ chemical transport models). As a result, graduates may have strong academic knowledge but lack the hands-on skills required to set up, run, and interpret real-world air quality situations. **These disconnects between theory and practice create a knowledge gap that limits the ability of local professionals to effectively address air quality challenges in their regions**.

Benchmarking examples of institutional housing of modeling systems: GEOS-chem is a global chemical transport modeling system at developed and maintained at Harvard University and GAINS is an integrated air quality management system developed and maintained at IIASA, Austria. GAINS is also an example of a reduced complexity model, which bypasses the use of a chemical transport model to convert a gridded emissions inventory into a gridded concentration field which can used for impact and scenario analysis





Additionally, most advanced applications of air quality modeling and analysis are concentrated in the United States and Europe, where funding, infrastructure, and technical expertise are more readily available. In contrast, LMICs have seen limited adoption of these tools, even though they often face some of the most severe air pollution challenges (Garland et al., 2024). Bridging this gap requires focused efforts to build local capacity through practical training, open information platforms, partnerships, and accessible technologies.

Technical personnel (Barrier risk – high)

A major challenge in LMICs is the lack of personnel with advanced technical training in air quality modeling and environmental management. This includes required expertise in computer engineering, geospatial information systems (GIS), data analytics, information management, communications, and finally meteorology and atmospheric science.

Even when individuals receive high-level education abroad or through specialized programs, many do not return to their home countries or remain engaged in local initiatives. This "brain drain" leaves a critical shortage of skilled professionals capable of implementing and managing sophisticated air quality systems. To address this, there is a pressing need for programs that not only provide technical training but also offer financial and operational support to encourage these professionals to stay and work locally.

Creating competitive career opportunities, offering incentives, and building a supportive ecosystem can help retain talent and ensure that local expertise is available to tackle air pollution challenges where it is needed most.

6. Way Forward Recommendations

The potential for air pollution research studies in India is immense, and we aim to highlight four key pillars that are essential for advancing this field, particularly in supporting the use of chemical transport models.

1	2	3	4
Enhancing data availability and quality is crucial; this includes improving emission inventories, meteorological data, and real-time monitoring systems to provide a solid foundation for modeling efforts.	Fostering targeted operational training programs can equip researchers and practitioners with the necessary skills to utilize chemical transport models; ensuring that they are adept at navigating the complexities of these tools.	Promoting interdisciplinary collaboration among scientists, policymakers, and stakeholders will facilitate the integration of diverse expertise, driving more holistic approaches to air quality management.	Increasing public engagement and awareness is vital; by involving communities in expanding the research efforts and encourage collective action.

Enhancing ground monitoring efforts will always be the central focus of air quality management, as it provides the critical foundation for regulations, audits, and progress reports. Reliable ground monitoring data is essential for understanding local air quality dynamics and assessing compliance with air quality standards. This data can be obtained through established methods, such as ground-based sensors, as well as from advanced satellite feeds that offer broader spatial coverage. By integrating both approaches, we can improve the accuracy and comprehensiveness of air quality assessments, ultimately supporting more modeling studies, regulatory frameworks and evaluations of targeted interventions.

On one side, the need for operational training is crucial and will also remain a priority for at least the next 10-20 years, institutionalizing the known chemical transport models in India and building a new crop of researchers and practitioners, who operate these models without fear of data.

The other side of the challenge is building localized emission inventories and consolidation of these inventories at the national and urban scales. At the time of this working paper, India still does not host an official consolidated emissions inventory to support air pollution modeling at any scale (national and urban airsheds). Here are some example intercomparison studies, which led to consolidation of inventories, establishing a baseline and modeling framework for the future modeling groups to follow.

1. Atmospheric Model Intercomparison Project (AMIP): https://www.wcrp-climate.org/modelling-wgcm-mip-catalogue/modelling-wgcm-mips-2/240-modelling-wgcm-catalogue-amip This program was initiated in the 1990s

to intercompare atmospheric models for use in climate research, including models that simulate air quality and atmospheric chemistry.

2. The Air Quality Model Evaluation International Initiative (AQMEII): https://joint-research-centre.ec.europa.eu/scientific-tools-databases/ensemble-atmospheric-chemistry-transport-and-dispersion-models/ensemble-case-studies/ensemble-air-quality-model-evaluation-international-initiative-aqmeii_en This program was initiated in the early 2000s to intercompare air quality models to improve their accuracy and provide better information for policymakers and the public.
3. The Hemispheric Transport of Air Pollution (HTAP): https://htap.org This program was initiated in the early 2000s to intercompare air quality models to improve understanding of the long-range transport of air pollutants across different regions of the world and build an ensemble emissions inventory for all the models.
4. The Model Inter-Comparison Study for Asia (MICS-Asia): https://www.acap.asia/en/research-main/mics-asia This program was initiated in the 2010s to intercompare emission inventories and air quality models specifically for the Asian region, where air pollution is a major public health concern.

For India and the Indian Subcontinent, an emissions and pollution modeling intercomparison exercise is due.

7. References

Crippa, M., Guizzardi, D., Butler, T., Keating, T., Wu, R., Kaminski, J., Kuenen, J., Kurokawa, J., Chatani, S., and Morikawa, T.: The HTAP_v3 emission mosaic: merging regional and global monthly emissions (2000–2018) to support air quality modelling and policies, Earth Syst. Sci. Data, 15, 2667–2694, https://doi.org/10.5194/essd-15-2667-2023, 2023.

Crippa, M., Guizzardi, D., Pagani, F., Schiavina, M., Melchiorri, M., Pisoni, E., Graziosi, F., Muntean, M., Maes, J., and Dijkstra, L.: Insights into the spatial distribution of global, national, and subnational greenhouse gas emissions in the Emissions Database for Global Atmospheric Research (EDGAR v8. 0), Earth System Science Data, 16, 2811-2830, https://doi.org/10.5194/essd-16-2811-2024, 2024.

Fowler, D., Brimblecombe, P., Burrows, J., Heal, M. R., Grennfelt, P., Stevenson, D. S., Jowett, A., Nemitz, E., Coyle, M., Liu, X., Chang, Y., Fuller, G. W., Sutton, M. A., Klimont, Z., Unsworth, M. H., and Vieno, M.: A chronology of global air quality, Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences, 378, 20190314, https://doi.org/10.1098/rsta.2019.0314, 2020.

Gani, S., Pant, P., Sarkar, S., Sharma, N., Dey, S., Guttikunda, S. K., AchutaRao, K. M., Nygard, J., and Sagar, A. D.: Systematizing the approach to air quality measurement and analysis in low and middle income countries, Environmental Research Letters, 17, 021004, https://doi.org/10.1088/1748-9326/ac4a9e, 2022.

Garland, R. M., Altieri, K. E., Dawidowski, L., Gallardo, L., Mbandi, A., Rojas, N. Y., and Touré, N. d. E.: Opinion: Strengthening research in the Global South–atmospheric science opportunities in South America and Africa, Atmospheric Chemistry and Physics, 24, 5757-5764, https://doi.org/10.5194/acp-24-5757-2024, 2024.

Guttikunda, S. K., Goel, R., and Pant, P.: Nature of air pollution, emission sources, and management in the Indian cities, Atmospheric Environment, 95, 501-510, https://doi.org/10.1016/j.atmosenv.2014.07.006, 2014.

Guttikunda, S. K., Nishadh, K. A., and Jawahar, P.: Air pollution knowledge assessments (APnA) for 20 Indian cities, Urban Climate, 27, 124-141, https://doi.org/10.1016/j.uclim.2018.11.005, 2019.

Guttikunda, S. K., Dammalapati, S. K., Pradhan, G., Krishna, B., Jethwa, H. T., and Jawahar, P.: What Is Polluting Delhi's Air? A Review from 1990 to 2022, Sustainability, 15, 4209, https://doi.org/10.3390/su15054209, 2023.

Holloway, T., Miller, D., Anenberg, S., Diao, M., Duncan, B., Fiore, A. M., Henze, D. K., Hess, J., Kinney, P. L., Liu, Y., Neu, J. L., O'Neill, S. M., Odman, M. T., Pierce, R. B., Russell, A. G., Tong, D., West, J. J., and Zondlo, M. A.: Satellite Monitoring for Air Quality and Health, Annual Review of Biomedical Data Science, 4, 417-447, https://doi.org/10.1146/annurev-biodatasci-110920-093120, 2021.

Liang, L., Daniels, J., Bailey, C., Hu, L., Phillips, R., and South, J.: Integrating low-cost sensor monitoring, satellite mapping, and geospatial artificial intelligence for intra-urban air pollution predictions, Environmental Pollution, 331, 121832, https://doi.org/10.1016/j.envpol.2023.121832, 2023.

Morawska, L., Thai, P. K., Liu, X., Asumadu-Sakyi, A., Ayoko, G., Bartonova, A., Bedini, A., Chai, F., Christensen, B., Dunbabin, M., Gao, J., Hagler, G. S. W., Jayaratne, R., Kumar, P., Lau, A. K. H., Louie, P. K. K., Mazaheri, M., Ning, Z., Motta, N., Mullins, B., Rahman, M. M., Ristovski, Z., Shafiei, M., Tjondronegoro, D., Westerdahl, D., and Williams, R.: Applications of low-cost sensing technologies for air quality monitoring and exposure assessment: How far have they gone?, Environment International, 116, 286-299, https://doi.org/10.1016/j.envint.2018.04.018, 2018.

Thunis, P., Kuenen, J., Pisoni, E., Bessagnet, B., Banja, M., Gawuc, L., Szymankiewicz, K., Guizardi, D., Crippa, M., Lopez-Aparicio, S., Guevara, M., De Meij, A., Schindlbacher, S., and Clappier, A.: Emission ensemble approach to improve the development of multi-scale emission inventories, Geosci. Model Dev., 17, 3631-3643, https://doi.org/10.5194/gmd-17-3631-2024, 2024.

www.urbanemissions.info