(Submitted version of the paper)

Comprehensive Air Quality Management System for Rapidly Growing Cities in Developing Countries

A. Kakarla*, M. V. Satish Reddy*, A. Qureshi †, S. Thatikonda†, S. De‡, S. Jana*
* Department of Electrical Engineering, Indian Institute of Technology Hyderabad
† Department of Civil Engineering, Indian Institute of Technology Hyderabad
‡ Department of Electrical Engineering, Indian Institute of Technology Delhi
*{ee13p1002,ee18resch01003,jana}@iith.ac.in, †{asif,shashidhar}@iith.ac.in
‡ swadesd@ee.iitd.ac.in

***Abstract*—In recent years, rapid industrialization and increased motor vehicle usage have made several developing cities unsustainable in terms of air quality among other factors. While air quality is monitored in such cities, monitoring stations are geographically separated, and the resulting measurements are often insufficient to infer pollution levels in requisite details. In addition, air quality information is in-comprehensively communicated to concerned groups leading to inefficient law enforcement. To this end, we envisage a system which initially makes extensive measurements from both the monitoring stations and densely-deployed low-cost compact sensors and later fuses the collected data with suitable modelling to generate detailed pollution maps. Finally, the system should disseminate the maps to the relevant stakeholders with necessary but sufficient abstractions. We believe such systematic architectures, when implemented at city-scale, will guide the governments towards transparent law enforcement, and firms towards maintaining a balance between economic development and environment friendly operations.**

***Index Terms*—air quality, health, policy making, smart city.**

I. Introduction

Since last several decades, rapid industrialization has been accompanying the bulging urban settlements, especially, in developing regions [1]. While industries arise out of economic necessities, those are often operated without any consideration for long-term sustainability. At the same time, increasing population leads to increase in the number of private and other vehicles, contributing to deteriorating air quality. Thus several cities, especially in developing countries, have extremely poor quality of air, and attendant respiratory and other medical ailments [2]. In response, certain efforts are being made in monitoring air quality around urban areas. However, such monitoring is most prevalent in and around developed countries, where industrial production generally takes place at large scales at isolated locations. More difficult challenges arise in developing cities. Specifically, as cities expand, highly polluted areas like industrial clusters and vehicular traffic hotspots often end up adjoining residential areas. To complicate matters, violation of emission norms is frequent (in fact, most violators pay scant regard for such norms), and such cities often lack effective monitoring tools [3]. Further, relevant data are often not available in suitable form for policy makers to fully comprehend and appreciate the situation on the ground. Moreover, common residents also do not have pollution information regarding their neighborhoods, and their health risks. In these circumstances, it has become imperative to develop a comprehensive air quality management system that not only gathers pollution information at various scales, but also systematically disseminates it to the stakeholders according to their needs.

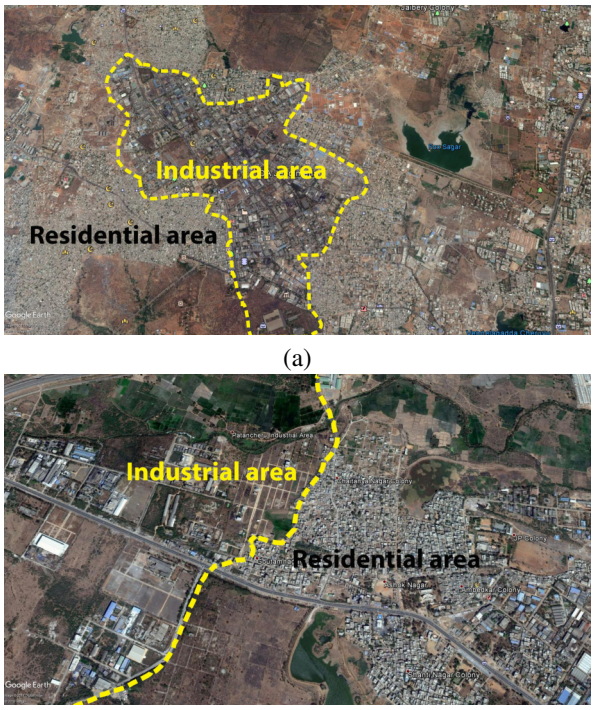

Fig. 1. Industrial proximity with residential areas in Hyderabad

Traditionally, to analyze atmospheric air quality, ambient air is sampled continuously at geographically distant but precise monitoring stations. [4]. Further, the collected data of different parameters is mapped to a time averaged air quality index (metric) based on exposure standards to aggregate and quantize the impact of the pollutants. These location wise metrics have inherent assumptions that the area in the vicinity of the monitoring station will report the same measurements and

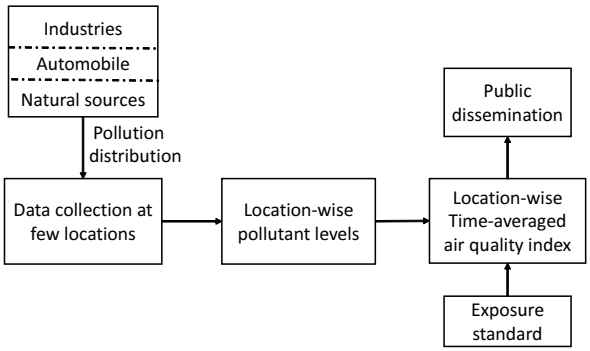

Fig. 2. Traditional air quality assessment

hence are also considered as spatially averaged metrics. Finally the metrics are disseminated to the citizens through public platforms as shown in Figure. 2.

In the short run, individuals of means may choose to stay in safer neighbourhoods based on this information. However, the long-term goal still remains to actually reduce pollution. To achieve such improvement, one should monitor the sources of pollution and regulate their emissions to be within permissible limits [5]. Especially, monitoring agencies can analyze the location-wise trends in the air quality metrics and mandate appropriate emission limits and enforce the firms to be follow the set standards (Figure. 3). Though such schemes are already put in practise, they still fall short of reducing the concerns about poor air quality possibly due to the following reasons. Firstly, the data collected by distant monitoring stations can provide the level of detail necessary to trace the emission violations. Additionally, a spatially averaged air quality metrics cannot be correlated with the health exposure studies and affect the efficiency of emission limits. Secondly, as the pollutants emitted from sources are affected by meteorological conditions, certain regions can be exposed frequently to the pollutants and still be undetected by the monitoring stations. Thirdly, the air quality information is not properly disseminated to all the stakeholders for them to make an informed decision.

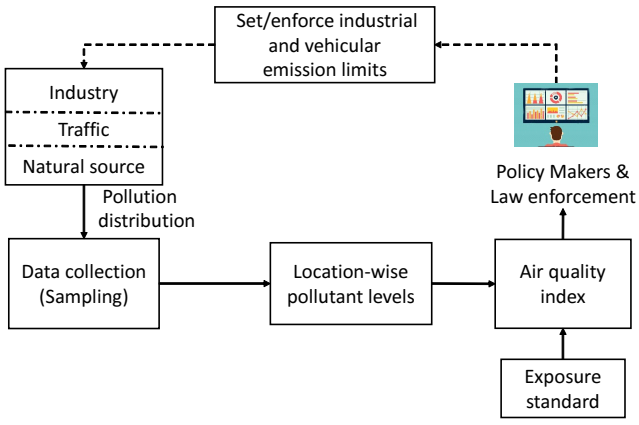

Fig. 3. Governed air quality management.

To address these issues, we need to develop an optimized management system in terms of sensing, modelling and data assimilation. In this paper, we highlight the areas in the conventional air quality monitoring system which should be optimized to address the aforementioned issues and envisage a comprehensive system level architecture. The rest of the paper is organized as follows. Section II presents the System architecture, while Section III describes the data collection and Section IV details further modelling. Risk assessment and data dissemination are detailed in Sections **??** and V respectively while Section VI concludes the paper.

II. Envisaged System Architecture

As mentioned earlier we identify the components of traditional air quality monitoring that should be optimized and further propose an improvised end-to-end system. Any air quality monitoring and management system is mainly be controlled by scientists who have the better hold of the technology. They should generate high-resolution pollution maps and disseminate it to address the concerns of stakeholders. To generate such maps, one should consider upgrading the traditional sensing mechanism. With the recent advancements in the sensing technology, inexpensive compact sensors can be the densely deployed alongside the existing monitoring stations near pollution hotspots to generate detailed air quality information. Deployment should include both static and mobile sensing mechanisms to acheive both spatial and temporal coverages. However, powering the sensors with such deployment is challenging due to (i) sensors will be mounted at unconventional locations and (ii) sensors are power hungry. So powering them using energy harvesting techniques ensures their continous operation. Next continously recording measurements should be sent to the server through respective communication technologies depending on the sensing strategy. For instance, static sensors should be comminucated through GSM technology due to their remote locations and mobile sensors through wifi/zigbee technologies. Further, the readings are post processed to reduce measurement artifacts such as cross-sensitivity and noise.

Though the idea is to place the sensors in multiple locations, deploying them every where is economically unfeasible. As a remedy, modelling the pollutant's environment fate and fusing it with the collected data can push towards higher spatio-temporal resolution with better accuracy [6]. In this context, many Eulerian models developed in the literature which are applicable for urban settings can be used. Note that any model is driven by both pollution source parameters and meteorology [7], [8] and success of any reliable model model relies depends on them. Additionally, all the traditional models have a grid nesting mechanism where the pollution maps are generated at different resolutions simultaneously to create multi-resolution spatio-temporal maps. Next, such maps should be remapped to air quality indices based on exposure standards. SUbsequently, disseminating the air quality information to all the stakeholders with necessary and sufficient level of detail is crucial. There are 4 main groups of stake holders: the general public, service

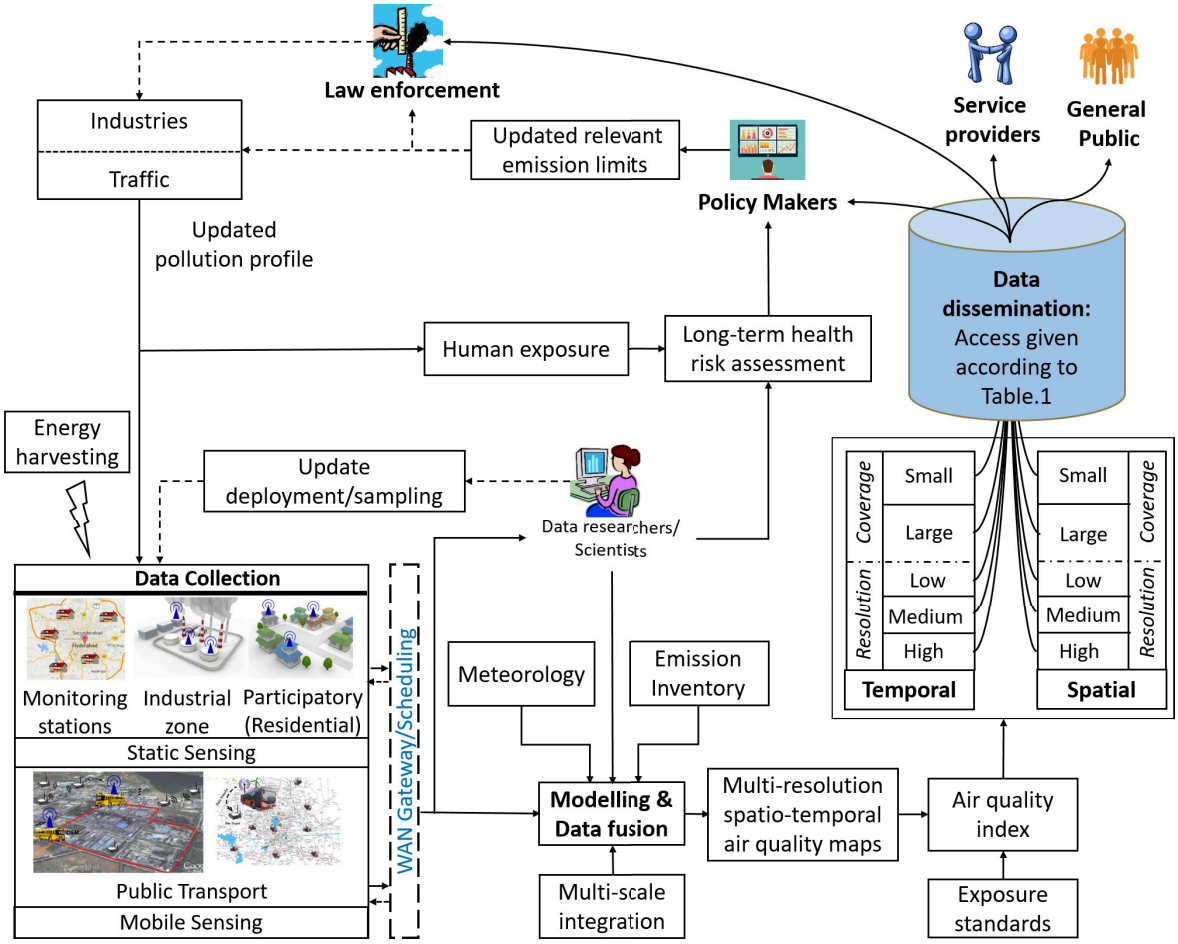

Fig. 4. Scalable Air Quality Management System Architecture

providers, policy makers and law enforcement. We propose a level of access that should be given to these stakeholders in Table.1. Public generally require local data with high resolution for longer duration while service providers need data for large spatial coverage with medium resolution. The policy makers look at larger coverage across both space and time, formulate and update policies based on the difference between recorded human exposure and the set exposure limits (standards). The regulatory authorities then enforce the updated limits on the anthropogenic sources improving the air quality. Thus, air pollution can be decreased by collecting detailed air quality information and systematically disseminating it. Now each of the individual component's improvements are further discussed in detail.

TABLE I

Stake holder	Temporal		Spatial	
	Resolution	Coverage	Resolution	Coverage
Public	Low	Large	High	Small
Service provider	Medium	Small	Medium	Large
Policy maker	Low	Large	Low	Large
Regulating authority	High	Small	High	Large

III. Data Collection

In a city, air can be polluted by many sources and each of their impacts are variable at different scales. The ultimate aim of air quality monitoring is to identify these sources and restrict them, if impossible to cease them completely. To this extent, the conventional air quality methodology is to acquire samples from geographically distant and highly accurate equipment. Further, this coarse-grained raw data are location-wise mapped to suitable metrics and disseminated to the public. However, such a methodology cannot detect certain sources distant from the monitoring stations and thereby cannot put down the concerns of people about the safety of their neighbourhoods. To overcome such difficulty, deploying the inexpensive reasonably accurate sensors along with the existing equipment through out the city which in turn will increase the monitoring costs. Accordingly, one needs to optimally deploy the sensors around the city depending on the segregation of sources.

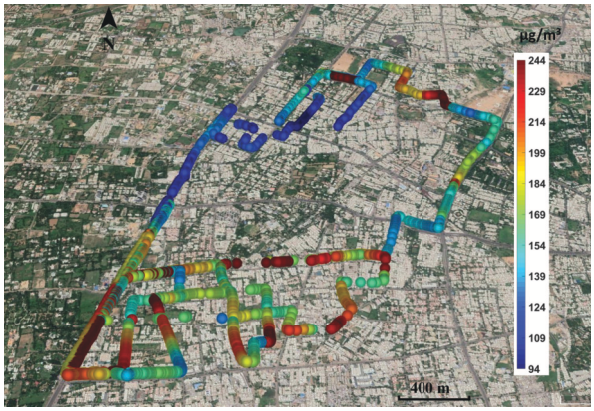

Fig. 5. Measured PM2.5 concentrations using mobile sensing strategy in a commercial area in Ahemadabad, India

A. Deployment Strategy

Sensors can be deployed with two different strategies.

1) Mobile sensing: : Measuring the air quality parameters using mobile vehicles is the most economic solution as huge coverage can be achieved with limited number of sensors. Specifically, as the roads ways and local rail tracks run throughout the length of the city and mounting the sensors on the public transport like city buses and local trains to collect data will give better spatial coverage and resolution [6]. See Figure. 5 for instance the mobile sampling of PM2.5 using a low cost sensor around a commercial area in Ahmedabad, India. However, as pollution is relatively high on roads and local train coverage is relatively low, profiles estimnated solely from mobile sensing tend to misrepresent the city's actual air quality level. Further, they lack temporal coverage which is crucial in the following scenarios. Firstly, abnormal emissions are often reported at industrial clusters and its difficult to isolate the actual violators. Secondly, high pollution levels are reported some times due to long range transport of pollutants from outside the city. If such events are not considered during policy making, sources might be over regulated. In each of these cases, as the temporal coverage is important, we also opt for static sensing.

2) Static sensing: : Static sensors should be deployed especially in voids created by mobile sensing and other pollution hot spots. At the same time, one have to consider historical data in different fields and obtain prior information about the safety levels through out the city to deploy sensors. This process is depicted though a flowchart in Figure. 6. Of course, employing such diverse sensing strategies also requires an efficient platforms to collect data and stored them. Note that each of these sensing mechanisms should have different internet gateways due to their inherent power and signal range requirements. Finally, encouraging a participatory approach of sensing makes the citizens also involved in the system. Precisely, concerned living communities should be encouraged to install the designed compact monitoring stations in their surroundings which continuously can sense and send the data.

B. Important Considerations

The use of compact low-cost sensor devices poses certain limitations including requirement of frequent calibration, inadequate minimum detection limits, sensitivity to temperature and humidity variations, cross sensitivity towards other gases and high power consumption. To overcome those, significant effort has been made. For instance, multiple linear regression has been suggested to model sensor sensitivity in various conditions along with temporal variation based on co-located sensor placement [9]. Further, artificial neural networks and regression analysis have been suggested to mitigate cross sensitivities and improve correlation [10]. Turning to power consumption, harvesting renewable resources like solar energy for these tiny instruments potentially makes them sustainable. It should be noted that the augmenting the monitoring stations with these compact sensors could mitigate various aforementioned drawbacks. Finally, it is also important to design a setup that is robust against pressure differentials caused during mobile sensing to which sensors are very sensitive.

IV. Air Quality Modelling

To estimate high resolution spatio-temporal maps from the recorded measurements, many statistical approaches like kriging interpolation were used in the literature [11]. However, such approaches lack proper scientific inference as they do not consider the physical phenomenon underlying the evolution of pollution profiles. For instance, certain sources could have more impact, even if their emissions are below the limit, because of the local meteorology. To be more precise, a pollutant when released from the source is transported by the wind and dispersed according to the atmospheric stability and urban conditions. Ideally, modeling this behaviour precisely and incorporating it in the profile estimation not

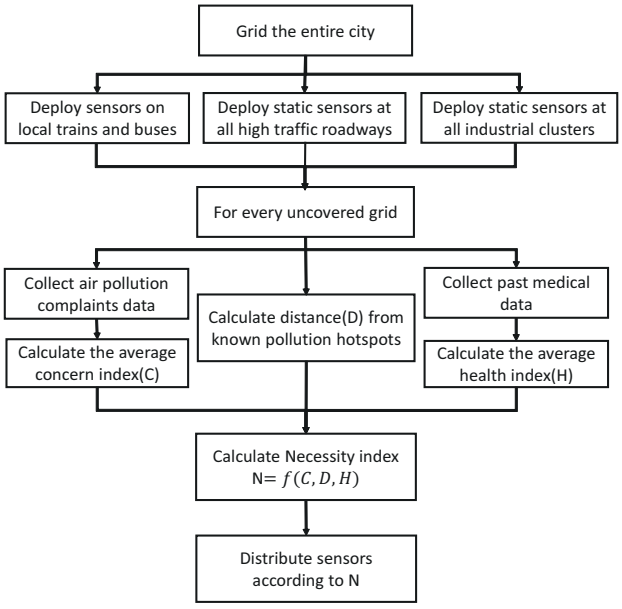

Fig. 6. Methodology for selecting static sensor locations.

only provides accurate pollution maps but also reduces the accuracy requirements of the sensing hardware. In practice, modelling mismatches tend to exist due to uncertainties in both meteorology and emission data. However, such mismatches are still less compared to interpolation techniques.

A. Choosing the right model

Many air quality models exist in the literature which have already been applied in many developing countries. However, each of these models have their own resolution and spatial coverage and choosing the the right model requires identifying the representative phenomenon in the city. For instance, in cities with tall structures, because of the air flow obstruction, turbulent dispersion is significant. To model such phenomena, computational fluid dynamics (CFD) models should be used[12]. On the other hand, cities with many pollution sources, but nominal building heights, should be dealt with Eulerian (or Lagrangian) models like EPISODE, WRF-Chem [13], [14].

B. Emission Inventory

Any model is driven by pollutant sources and the meteorology. Many a times, biggest challenge for the researchers is to identify these sources and estimate their emission rate. Moreover, acquiring such data on a city scale is very difficult. Now, different challenges in dealing with anthropogenic sources of pollution mainly i.e. industries and vehicular traffic using densely deployed sensors is detailed now.

1) Industrial sources: One should collect the emission data of all point and non-point industrial sources. Recently, some of the developing countries have mandated industries to report their emission instantaneously [15]. However, emission violations can still occur if the enforcement lacks detailed air quality information especially in the industrial vicinity. Moreover, as long as the measuring devices are within the industries, they are prone to be tampered. In other words, the emission inventory, if estimated from within-industry records, many inaccuracies and mismatches arise. To address this issue, the idea of deploying the compact sensors around the industrial cluster provides more authentic data. The deployment strategy based on Figure. 6 gives sufficient samples around these pollution spots. Accordingly, from this data, they can also use source localization tools to get a prior information about the location and the emission levels of the violators [16], [17].

2) Vehicular traffic: Conventionally, to estimate the emission values from vehicular traffic, the extent and the volume of automobile usage is mapped to corresponding emission values based on already estimated per unit emission data. However, such data is not only limited, but is also bound to errors due to their dynamic emission characteristics while running is challenging to model. Thus, rather than relying on a particular vehicle emission, considering emissions from a whole road lane as an emission source reduces disparities in modelling. Such road level emissions only be captured by both static and mobile sensors around the road lanes.

C. Data fusion

The air quality models predict the atmospheric state based on the meteorological and emissions data. As mentioned earlier, modelling errors tend to exist along with measurement artifacts. So, synergizing the information from both measurement and model is the best way of estimating the pollutant profile. The state-of-the-art models use a data fusion technique called 4D-Var which is a a Maximum a Posteriori (MAP) estimate.

D. Multi-scale integration

To give a very detailed resolution around pollution hot spots, many of the models like WRF, FLEXPART incorporate grid nesting mechanism [14]. In other words, multiple grids are defined with different resolution and each of them are resolved at every instance. The interaction between coarser and finer grids is also solved. However, simulating such mechanism for every instance throughout the year is infeasible due to computational complexity and hence the necessary hardware requirements, especially for developing countries. To complicate matters, there is no suitable mechanism that directly fuses nested grids with densely deployed sensor measurements. In this regard, a recent study, although demonstrated their method on temporally averaged data, have integrated coarser grid with finer grids using universal kriging [13].

E. Safety Metrics

Communicating the analytics and the information estimated from the collected readings to the general public in a efficient and useful manner is important. Urban air pollution comprises of mixture of several pollutants emitted from different anthropological activities. Individual safety metrics for each of the pollutant not only complicates the decision making for policy makers, but also are incomprehensible for the general public. Further, it should also evaluate trade of involved in different control policies. For instance, replacing coal with compressible natural gas decreases carbon dioxide at the cost of increased methane. An air quality index (AQI) is one such measure defined as single number for reporting the air quality with respect to different pollutants on the human health. Further, it also enables one to compare overall air pollution in different localities. All toxics and particulate are finally mapped to air quality indices shown in Table. II).

F. Long-term health risk assessment

Long-term risk assessment aims to estimate the risks of past, current or future exposure to air pollution and of changes in exposure that may result from planned policies. Such assessment is based on relationship between a exposure and incidence of a health effect. The extent of exposure is obtained from the generated pollution maps and the later from recorded human exposure data accessed through proper collaboration with medical facilities. Generally, the hospitals have records of the patients health along with his/her residence information which can be used to create average health maps and further be correlated with the pollution maps.

TABLE II
AIR QUALITY INDICES

Air quality category	Range
Good	0-50
Satisfactory	51-100
Moderate	101-200
Poor	201-300
Very poor	301-400
Severe	401-500

V. DATA DISSEMINATION

Once the spatio temporal maps are estimated, they should be provided to all stakeholders. The main stake holders are general public who are directly exposed to the pollution, the service providers who use the air quality data, policy makers who have to ensure the citizen well being, and regulatory authorities who should keep the source emissions under control. Now we detail our proposed scheme of giving data access to all the stake holders. The level of access is given in Table.I will be justified in this section.

General Public. Generally, every person desires his/her localities to be sustainable over last few years with less spatial details. Hence data with low temporal and low spatial resolution for longer duration should provided to general public.

Service Providers Service providers are the groups which provide services like pollution-free travel routes suggestions. They use the air quality data in their algorithms to provide routes which are safer. They should be given medium resolution of spatio-temporal data for the whole city for short duration every time[18].

Policy makers Policy Makers should make policies that guide cities towards sustainability. They analyze the source emissions and the trade offs between the development of the city and the average health of citizens due to pollution exposure and set optimal standards. Thus, data of whole city for last few years at low resolution should be provided to policy makers.

Law enforcement Law enforcement also termed as regulating authorities monitor the emission levels of all the sources including industrial and vehicular emissions. They ensure that all the firms comply with the set emission guidelines. To this end, they need high resolution information for a short duration.

VI. DISCUSSION

With the sudden increase in the ubiquotious networks, the Smart Cities concept has evolved and is already drawing the governments and businesses into it. Many cities are already embracing the information and communication technology towards transparent urban governance and productive, sustainable habitats. However, planning a smart city in developing countries should address the issues detailed in section 1. To this end, the optimized architecture described in this paper mitigates pollution to great extent. The dense deployment included in the architecture enhances the detail of a varying pollution profile of a dynamic city environment. We acknowledge that though state-of-the art of air quality monitoring has been demonstrated long before, implementing them at systemic governance levels need sufficient reliability which is expected to be met in future. For such futuristic implementation, data flow between all the components in the architecture should be well defined which was the main focus of this paper. Finally, the data dissemination scheme proposed in this paper helps all the stakeholders to make informed decisions. For instance, along with displaying the safety indices openly in the alarming zones, providing analytic about the sources like vehicles daily emissions could enhance awareness in the general public. Such practises motivates the public to strive for a healthy environment.

ACKNOWLEDGMENT

This study was partially funded by Department of Electronics and Information Technology (DEITY), India (13(2)/2012-CC&BT) and JST/JICA SATREPS, Japan.

REFERENCES

[1] Y. Wang, L. Yang, S. Han, C. Li, and T. Ramachandra, "Urban co 2 emissions in xian and bangalore by commuters: implications for controlling urban transportation carbon dioxide emissions in developing countries," *Mitigation and Adaptation Strategies for Global Change*, vol. 22, no. 7, pp. 993–1019, 2017.

[2] "Air pollution," https://www.who.int/airpollution/en/.

[3] G. McGranahan and F. Murray, *Air pollution and health in rapidly developing countries*. Earthscan, 2012.

[4] K. Clemitshaw, "A review of instrumentation and measurement techniques for ground-based and airborne field studies of gas-phase tropospheric chemistry," *Critical Reviews in Environmental Science and Technology*, vol. 34, no. 1, pp. 1–108, 2004.

[5] A. Jordan and E. Matt, "Designing policies that intentionally stick: policy feedback in a changing climate," *Policy Sciences*, vol. 47, no. 3, pp. 227–247, 2014.

[6] D. Hasenfratz, O. Saukh, C. Walser, C. Hueglin, M. Fierz, and L. Thiele, "Pushing the spatio-temporal resolution limit of urban air pollution maps," in *2014 IEEE International Conference on Pervasive Computing and Communications (PerCom)*, March 2014, pp. 69–77.

[7] S. Madala, K. H. Prasad, C. Srinivas, and A. Satyanarayana, "Air quality simulation of no x over the tropical coastal city chennai in southern india with flexpart-wrf," *Atmospheric Environment*, vol. 128, pp. 65–81, 2016.

[8] P. Holnicki, A. Kałuszko, and W. Trapp, "An urban scale application and validation of the calpuff model," *Atmospheric Pollution Research*, vol. 7, no. 3, pp. 393–402, 2016.

[9] B. Maag, O. Saukh, D. Hasenfratz, and L. Thiele, "Pre-deployment testing, augmentation and calibration of cross-sensitive sensors." in *EWSN*, 2016, pp. 169–180.

[10] L. Spinelle, M. Gerboles, M. G. Villani, M. Aleixandre, and F. Bonavitacola, "Field calibration of a cluster of low-cost available sensors for air quality monitoring. part a: Ozone and nitrogen dioxide," *Sensors and Actuators B: Chemical*, vol. 215, pp. 249–257, 2015.

[11] Y. Ramos, B. St-Onge, J.-P. Blanchet, and A. Smargiassi, "Spatio-temporal models to estimate daily concentrations of fine particulate matter in montreal: Kriging with external drift and inverse distance-weighted approaches," *Journal of Exposure Science and Environmental Epidemiology*, vol. 26, 12 2015.

[12] E. Rivas, J. L. Santiago, Y. Lechón, F. Martín, A. Ariño, J. J. Pons, and J. M. Santamaría, "Cfd modelling of air quality in pamplona city (spain): Assessment, stations spatial representativeness and health impacts valuation," *Science of the Total Environment*, vol. 649, pp. 1362–1380, 2019.

[13] P. Schneider, N. Castell, M. Vogt, F. R. Dauge, W. A. Lahoz, and A. Bartonova, "Mapping urban air quality in near real-time using observations from low-cost sensors and model information," *Environment International*, vol. 106, pp. 234 – 247, 2017.

[14] G. A. Grell, S. E. Peckham, R. Schmitz, S. A. McKeen, G. Frost, W. C. Skamarock, and B. Eder, "Fully coupled online chemistry within the wrf model," *Atmospheric Environment*, vol. 39, no. 37, pp. 6957–6975, 2005.

[15] "Online monitoring of industrial emission effluent," http://cpcb.nic.in/Online-Monitoring-Industrial-Emission-Effluent/.

[16] A. Kakarla, A. Qureshi, T. Shashidhar, S. De, S. G. Singh, and S. Jana, "Source localization via aermod-based simulation under mean squared error criterion: Demonstration using field data," in *IEEE International Geoscience and Remote Sensing Symposium*, July 2017, pp. 6201–6204.

[17] B. Hosseini and J. M. Stockie, "Bayesian estimation of airborne fugitive emissions using a gaussian plume model," *Atmospheric Environment*, vol. 141, pp. 122 – 138, 2016.

[18] "M2smart project," m2smart.org/en/.