An Improvement on Air Pollution Modelling using Mobile Sensors

Conference Paper · November 2015

2 authors:

Manoochehr Kheradmandi
University of Tehran

Rahim A. Abbaspour
University of Tehran

An Improvement on Air Pollution Modelling using Mobile Sensors

Kheradmandi, M. * and A.Abbaspour, R.

Department of Surveying Engineering, College of Engineering, University of Tehran, Tehran, Iran
(kheradmandi.m, abaspour)@ut.ac.ir

KEYWORDS: Data Fusion, Air pollution, Mobile sensing, Kriging.

ABSTRACT:

Air quality monitoring is extremely important as air pollution has a direct impact on human health. Planning to prevent and reduce air pollution requires a complete understanding of how the spatial distribution of pollutants, which in turn requires increasing the number and density of Air Quality Monitoring Networks (AQMNs). But the high cost of implementation and maintenance of these sensing stations is the most important obstacle to develop the network. On the other side, data fusion is one of the methods that make it possible to achieve more accurate, precise, and specific interferences by combining data from multiple sensors and related information. This paper illustrates integration of some datasets collected by mobile sensors simultaneously and Tehran's AQMNs. The results of integration of measured data in comparison to stationary stations measurements is so notable. RMSS -a standard metric to evaluate the accuracy of air quality models- on the integrated stationary and mobile sensor measurements is closer to 1 rather than measurements of only stationary stations. Significant results of this research shows that wide range use of mobile sensors and integrating mobile and stationary sensors data significantly improves real time pollution assessment.

1. INTRODUCTION:

Data fusion is one of the methods that make it possible to achieve more accurate, precise, and specific interferences, by combining data from multiple sensors and related information. Because of uncertainty and imperfectness in data from multiple sources, data fusion of multiple sensors is more efficient in comparison with using a single, independent sensor. Fused data from multiple sensors provides several advantages over data from a single sensor. Using the same multi-sensor data, statistically reduces the error, increases the robustness of the system and causes complete information. Sometimes, Integration of sensor data can be more economically effective (Hall, 2001, Liggins 2008). However the concept of data fusion is not new, the emergence of new sensors and advanced processing techniques have made fusion of data increasingly viable.

In addition, Air pollution has a strong influence on the quality of human beings life and air pollution it is the greatest problem in mega cities that controlling it can reduce mortality rates. Planning to prevent and reduce air pollution requires a complete understanding of how the spatial distribution of pollutants are, which in turn requires increasing the number and density of air quality monitoring networks. But the high cost of implementation and maintenance of these sensing stations is the most important obstacle to develop the Air Quality Monitoring Networks (AQMNs) (Devarakonda, 2013, Hasenfratz , 2012).

Multi sensor data integration for environmental applications such as air pollution monitoring is often desirable due to different types of measurements obtained by different sensor networks implemented at distinctive spatial and temporal scales.

The integration of satellite and ground-based sensing provides an example of how data can be combined at different spatiotemporal scales. One of the main deficiencies of Air Quality Monitoring Networks is their sparse spatial distribution. In contrast, geospatial data products derived from satellite borne sensors provide spatially explicit 'Big Data' on a variety of pollutants.

However, current air pollution measurements from satellite borne sensors also have significant limitations, including data gaps resulting from clouds, limited temporal resolution (Duncan, 2014). Therefore it seems that increasing the number and density of ground-based air quality monitoring sensors is a more efficient Strategy to achieve an optimal air pollution monitoring.

The capabilities and availability of cheaper, more sensitive and sophisticated sensors for gases, particulates, other environmental measurements have improved and are enabling researchers to collect data in unprecedented spatial, temporal and contextual detail (Stocker, 2014).

Using of low cost sensors probably leads to higher spatial resolution for environmental monitoring because of broader sensing with non-traditional researchers, environmental justice organizations and even citizen scientists to participate in collecting environmental.

Low cost gas sensors are small, portable and can be assembled on vehicles for collecting air pollution levels

* Corresponding author

almost everywhere, as a result large amount of data from different places are recorded.

Consequently, the advent of new sensor technology can be used not only in urban areas with large spatiotemporal variations in pollutant concentrations and but also in rural areas where few conventional Air Quality Monitoring station may be available.

While some types of sensors are already commonly presented in consumer devices (e.g., geolocation, motion, sound, etc.), and becoming ubiquitous in everyday life, generating data at an unprecedented rate and scale; other kinds of compact, high-power sensors (e.g., air quality sensors) are not yet commonly included but offer the ability to collect additional data of individual and social interest (Aoki, 2008).

Three basic alternatives can be used for multisensor data fusion:

1. direct fusion of sensor data
2. information fusion achieved of sensor data
3. inference or decision subsequently achieved of sensor data

Considering this point that outputs of Stationary and Mobile sensors are commensurate (i.e., if the sensors are measuring the same physical phenomena and same unit) direct fusion of sensor data can be used. In the rest of this paper after investigating the features of Mobile Sensor and study the area in the following, modelling spatial distribution of air pollution by integrated sensor data have been described.

2. MOBILE SENSOR UNIT

The Mobile Sensor Unit shown in Figure 1 includes two sensors for measuring carbon monoxide with metal oxide and electrochemical technologies, temperature and humidity sensor for accurate and reliable calibration and achieve much more accurate pollution levels in different temperature and climate conditions, Bluetooth to send information to mobile phones, SD card to store measured values and GPS for positioning designed unit has been used.

2.1. Calibration of Mobile Sensor Unit

Low-cost gas sensors must be frequently re-calibrated as they are unstable and responsive to the influence of interfering gases (Hasenfratz, 2012). In order to ensure data accuracy, carbon monoxide sensors was calibrated by using Analyzer 12M device, in addition temperature and humidity parameters were modeled by using linear regression to achieve better calibration.

High quality data is required to develop reliable maps of air pollution. Evaluating the quality of collected data in a large urban area for this purpose is a challenging task. So in order to assess the quality of sensor Measurements, comparison of Sensor Measurements with Static Air Quality Monitors has been done. Pearson correlation coefficient for the TarbiatModares and Setadbohran, respectively, 0.94 and 0.96 was obtained which indicates relatively high quality measurements of Mobile Sensors.

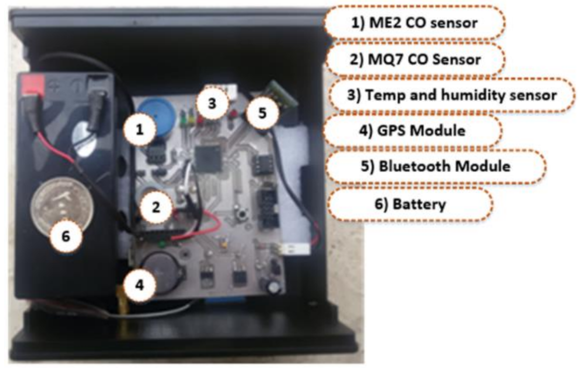

Figure 1. Mobile Sensor Unit for measuring Co emissions

2.2. Measuring Air Pollution through different path:

Because the main purpose of this paper is to examine how to use mobile sensors and integrate mobile sensor data with stationary air quality network to model the distribution of carbon monoxide measuring emissions of Co. in several streets of Tehran with different traffic volumes was performed that the Mobile sensor path on the satellite image of Tehran mapped and shown in figure2.

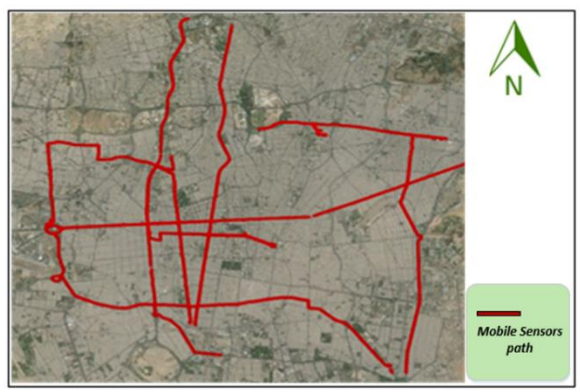

Figure 2. Mobile sensor path

2.3. Filtering Measurements of Mobile sensor unit

Ensuring a high data quality by calibrating and filtering measurements is required. Assessing and ensuring the quality of collected data is done in three following steps

1. Remove the GPS data with horizontal dilution of precision (HDOP) values above 3. The HDOP value specifies the GPS location's precision based on the geometric positioning of the GPS satellites. Values below 3 denote a good to excellent positioning within a few meters.

2. Remove the data collected in the early minutes of the recording.
3. Remove the data that show the bad performance of the system, such as too high or too low values with Grubbs method. Grubbs' test (Grubbs, 1969 and Stefansky, 1972) is used to detect a single outlier in a univariate data set that follows an approximately normal distribution. In total, we collected over 17000 measurements that after filtering of collected data only 16700 measurements of mobile sensor remained which in comparison with the lower number of fixed Air Quality monitoring air pollution stations is a significant number. Comparison with the lower number of fixed Air Quality monitoring air pollution stations is a significant number.

3. FUSION OF STATIONARY AND MOBILE SENSORS:

To fuse the data of static, stationary AQMNs with those from mobile sensors, a methodology is designed (Fig. 3) and developed. According to this architecture, two datasets are entered in a data integration module. This architecture is evaluated over the dataset of static stations of municipality of Tehran and some datasets collected by mobile sensors simultaneously. The results on the measured points show the priority of proposed method over the current modeling, which is based only on the static stations.

4. SPATIAL DISTRIBUTION OF AIR QUALITY:

Because of the most number and high density of Air Quality Monitoring Networks in Downtown of Tehran, this region is considered as the study area to assess air quality in this project. The Table 1 lists the stations name with their coordinates in UTM zone 39N. Spatial distribution of air pollutants with sufficient accuracy and high spatial resolution is the key piece of information for assessing the risks to human health or evaluating the air quality policy quantitatively. An Interpolation method can been used successfully for predicting the spatial distribution of the mean concentrations of air pollutants mainly for epidemiological studies (Ross, 2007)

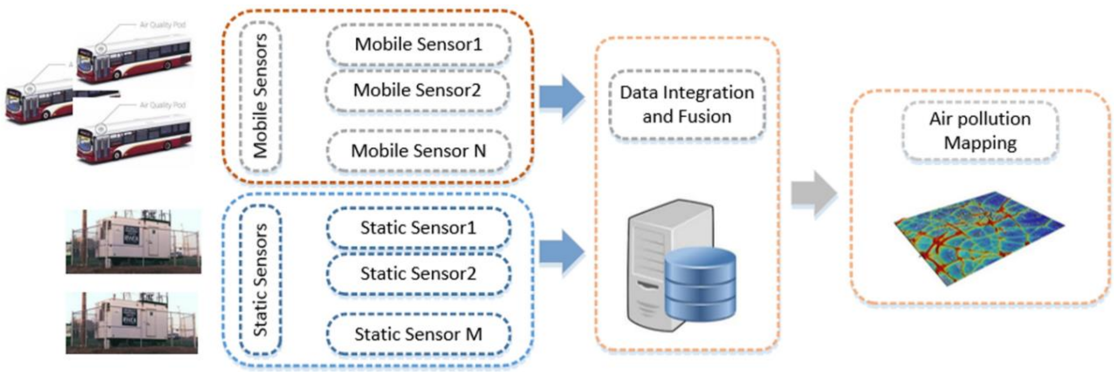

Figure 3: proposed methodology for integration of Mobile and Static Air Quality Sensors

Table1. Air quality monitoring stations name with their coordinates in UTM zone 39N

No.	Station Name	Northing(Meter)	Easting(Meter)	No.	Station Name	Northing(Meter)	Easting(Meter)
1	Tarbiat Moddares station	٣٩٥٨٥٣٠	٥٣٤٧٠٣	8	Daroos station	٣٩٥٢٩٨٠	٥٤١٠٥١
2	Sharif Uni. staion	٣٩٦١٨٨٠	٥٣١٢٨١	9	Aghdasieh Station	٣٩٥١٣٩٠	٥٤٣٨٢٥
3	Region 16 Municipility station	٣٩٥٩٢٩٠	٥٣٦٠٠١	١٠	Region 2 Municipility station	٣٩٤٤٦١٠	٥٣٣٢٧٦
4	Region 11 Municipility station	٣٩٥٠٤٥٤	٥٣٥٢٧٠	١١	Region 10 Municipility station	٣٩٤٧٧٥٠	٥٣٢٣٩١
5	Region 4 Municipility station	٣٩٥٣٧٦٠	٥٤٤٥٣١	١٢	Setad bohran station	٣٩٥٥٤٢٠	٥٣٨٩٩٦
6	Piroozi station	٣٩٥٧٦٩٦	٥٤٤٦٧٢	١٣	Poonal station	٣٩٥٠٣٤٠	٥٢٨٤١٥
7	Golbarg station	٣٩٤٨٣٧٨	٥٤٥٧٧١	١٤	Fat-h Station	٣٩٥٤٢٣٠	٥٣٠٥٤٤

.

There are different methods to assess spatial variation of air pollutants. Kriging have applied for the prediction of the spatial distribution of air pollutants in some studies (e.g., Mercer, 2011; Hengl, 2003). Simple kriging interpolation method was used for interpolation of carbon monoxide emissions on two sets of data. The first data set are the measurements that done on Air Quality Monitoring Networks stations and the second data set are integration of measurements of Mobile sensor and Stationary sensors. Interpolation results on these two data sets is shown below.

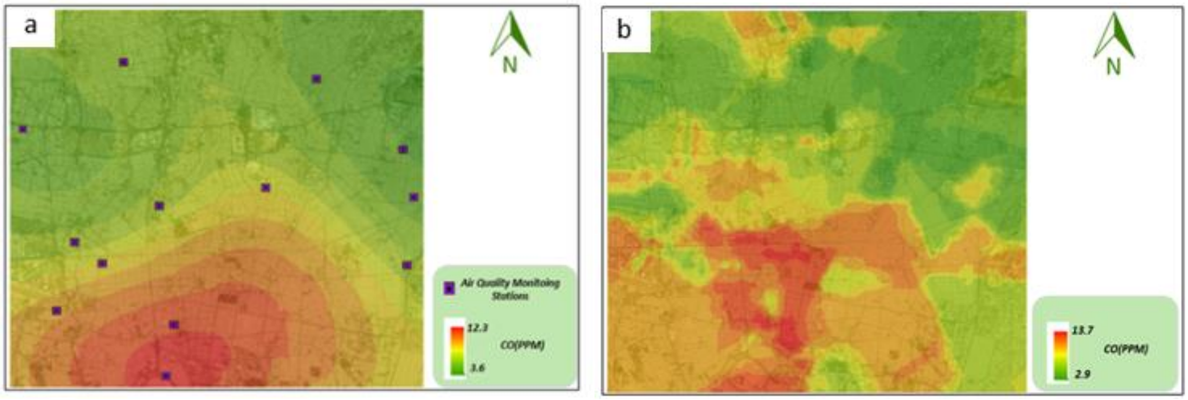

Figure 4: a) interpolation on measurements of Air quality monitoring stations. b) Interpolation on Integrated Mobile Sensor and air quality monitoring stations measurement

5. COMPARISON OF RESULTS

Optimality and Validity are two issues for assessing efficiency of interpolation results from different data sets. Smaller root mean squared prediction error for a particular model denotes more optimal model. Being closer the root-mean-squared prediction error to the average estimated prediction standard error for a particular model denotes more valid model because only the estimated standard errors can be assessed for predicting uncertainty. When the root-mean-square standardized is close to one and the average estimated prediction standard errors are close to the root-mean-squared prediction errors from cross-validation, the appropriation of the model is so obvious.

According to the mentioned points and comparison of the interpolation results shown in Table 2, Root-Mean-Squared-Standardized on the integrated stationary and Mobile sensor measurements are closer to 1 rather than Stationary Stations Measurements and also differences between root mean square and average standard error on the integrated stationary and Mobile sensor measurements are so smaller than differences between root mean square and average standard error on only stationary stations measurements. Another criteria that can be used to assess performance of interpolation on datasets is examining the relationship between the measured amounts and estimated values by the interpolation method. Relationship between the measured amounts and values estimated by the interpolation method for integrated stationary and Mobile sensor measurements and Stationary Stations Measurements are shown in Figure 4.

Table 2. Comparison of the interpolation results

Studied statistic	Stationary stations Measurements	Integrated stationary and Mobile sensors Measurements
Root-Mean-Square	1.3430	0.579
Root-Mean-Square Standardized	0.773	0.910
Average Standard Error	1.717	0.600

Another criterion that can be used to assess performance of interpolation on datasets is examining the relationship between the measured amounts and estimated values by the interpolation method. Relationship between the measured amounts and values estimated by the interpolation method for integrated stationary and Mobile sensor measurements and Stationary Stations Measurements are shown in Figure 5.

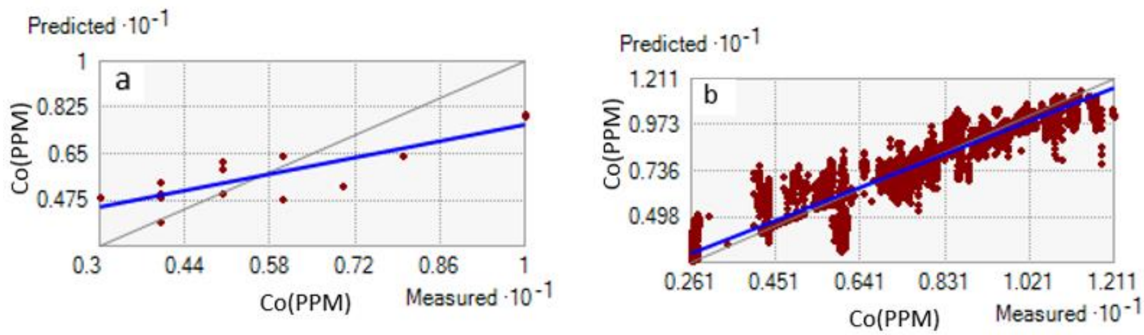

Figure5 (a): Relationship between the measured amounts and values estimated by the interpolation method for Stationary Stations Measurements. (b): Relationship between the measured amounts and estimated values by the interpolation method for integrated stationary and Mobile sensor measurements.

Closer to 1 for slope of fitted line in Figure4: (b) rather than Figure5: (a) indicates higher potential of integrated stationary and mobile sensors measurements to estimate air pollution emissions. The interpolation error for each of integrated stationary and Mobile sensor measurements and Stationary Stations Measurements were Calculated, mapped and shown on the following Figure (Fig. 6).

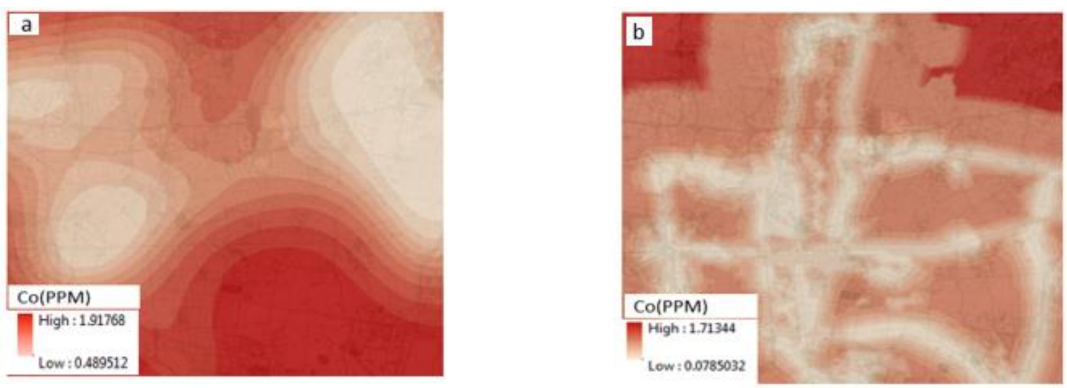

Figure 6. (a) Interpolation error for integrated stationary and Mobile sensor measurements, (b) Interpolation error for stationary sensor measurements.

6. CONCLUSION

Mobile Sensors are one interesting recent development in wireless and human sensing that they are extremely multifunctional and may be useful for different applications, such as environmental monitoring. On the other side Monitoring and controlling gas pollutant concentrations are necessary measurements to tackle urban atmospheric pollution as a crucial issue, and using mobile sensors can make a significant contribution to air pollution monitoring. But Low cost mobile sensors cannot accurately measure air pollution exposure. Reviewing what is mentioned in this article shows that using of mobile sensors and combining mobile data sets with data from fixed stations lead to increasing precision of air pollution maps. But in addition to imprecise measurements of mobile sensors, using them requires frequent calibration which makes it difficult to use them. Another challenge that most of authors also pointed out is battery life that compel users to recharge the device after several hours of use.

REFERENCES

P. M. Aoki1, 2008. Common sense: Mobile environmental sensing platforms to support community action and citizen science. In*: UbiComp 2008, Seoul, South Korea, September 2008*

Devarakonda, S., 2013. Real-Time Air Quality Monitoring Through Mobile Sensing in Metropolitan Areas. *Proceedings of the 2nd ACM SIGKDD International Workshop on Urban Computing (UrbComp'13), Chicago, USA, 11–14 August 2013*

Duncan, B.N., 2014. Satellite data of atmospheric pollution for U.S. air quality applications: examples of applications, summary of data end-user resources, answers to FAQs, and common mistakes to avoid. *Atmospheric Environment 94 (2014): pp. 647-662.*

Grubbs, F. E. 1969. Procedures for detecting outlying observations in samples. *Technometrics, 11(1), pp. 1-21*.

Hasenfratz, D. 2012. Participatory Air Pollution Monitoring Using Smartphones. In*: 2nd International Workshop on Mobile Sensing.*

Hall, D. L., 1997. An introduction to multisensor data fusion. *Proceedings of the IEEE, 85(1), pp. 6-23.*

Hengl, T. 2003. Comparison of kriging with external drift and regression-kriging. *Technical note, ITC, pp. 51*.

Liggins, M. E., 2008. Multisensor data fusion. CRC Press, Inc.

Mercer, L. D., 2011. Comparing universal kriging and land-use regression for predicting concentrations of gaseous oxides of nitrogen (NO x) for the Multi-Ethnic Study of Atherosclerosis and Air Pollution (MESA Air).*Atmospheric Environment, 45(26), pp. 4412-4420.*

Ross, Z., 2007. A land use regression for predicting fine particulate matter concentrations in the New York City region. *Atmospheric Environment, 41(11), pp. 2255-2269.*

Stefansky, W. 1972. Rejecting outliers in factorial designs. *Technometrics, 14(2), pp. 469-479.*

Stocker, M., 2014. Representing situational knowledge acquired from sensor data for atmospheric phenomena*. Environmental Modelling & Software, 58, pp. 27-47*.