Built Environmental Journal *Vol. 3, No. 2, 33-41, 2006*

MICRO-SCALE AIR QUALITY MONITORING USING REMOTE SENSING AND GIS TECHNOLOGY

Rohayu Haron Narashid and Wan Mohd Naim Wan Mohd

Department of Surveying Science and Geomatics, Faculty of Architecture, Planning and Surveying , Universiti Teknologi MARA, Shah Alam, MALAYSIA

ABSTRACT

Air quality monitoring is needed to control air pollution in urban areas. A limited number of air quality monitoring stations, limits the initial strategy of pollution prevention program in Malaysia in air quality monitoring especially at micro-scale level. The technique used to monitor air quality involves manual measurement of pollution concentrations within the area of measuring station. A more practical approach is needed to study air quality at micro-scale. One possibility is the use of satellite remote sensing and Geographical Information System (GIS) technologies. This paper present results obtained from research carried out to assess the suitability of satellite remote sensing and GIS techniques for air quality monitoring. Although three different datasets were used in the research, only results obtained from the Landsat 7 Enhance Thematic Mapper Plus (ETM+) dataset dated the 15th of July 2000 will be described. Eight (8) Continuous Air Quality Monitoring (CAQM) data are used to determine the relationship between digital number (DN) of the thermal infrared band and four pollutant parameters, i.e. Sulfur Dioxide (SO_2), Nitrogen Dioxide (NO_2), Carbon Monoxide (CO_2) and Particulate Matters (particles less than 10 micron in size – PM_{10}). The study area for this research covers the whole of Klang Valley Region. A limited area, which covers only part of the Shah Alam city is used for micro scale study. Since there are limited numbers of CAQM stations available, the concept of "Virtual Stations" is introduced to densify these stations. Based on derived pollutant values of these stations, kriging interpolation method is used to generate the pollutant maps of different pollutant parameters of the study area. Results from this research have indicated that SO_2, NO_2, CO and PM_{10} have positive correlation with digital number of the thermal infrared band. Initial analysis has shown that there is a strong correlation between satellite digital number (thermal band) and SO_2 pollutant. Maps of pollution concentration generated from virtual stations using kriging interpolation method produce a more realistic distribution of air pollution. The concentration of SO_2, NO_2, CO and PM_{10} are highest in the industrial zones of Shah Alam.

Keywords: Micro-Scale, Remote Sensing and GIS Technology

Introduction

Air pollution is one of the major environmental concern throughout the world today. Malaysia, as one of the most develop nation among the developing countries faces similar situation. Between 1992 and 1993, the Department of Environment (DOE) Malaysia and a team from the Japanese International Cooperation Agency (JICA) has carried out air quality study in the Klang Valley

ISSN 1675-5022

Region. Report this study *(Air Quality Management Study for Klang Valley Region)* and yearly air quality reports published by the DOE and a report published by Sham *et al*., (1991) have shown that traffic, industries and open burning contributed to a high level of particulates.
An hourly measurement of carbon monoxide, sulfur dioxide and nitrogen dioxide at the Kuala Lumpur City Hall, Petaling Jaya and Shah Alam have shown that levels of these pollutants exceeded the recommended Malaysian Guidelines.

According to Stern *et al.* (1984), air pollutants can be classified into primary and secondary pollutants. Primary pollutants are emitted directly from sources into atmosphere. However, the primary pollutants do not, of themselves, produce all of the adverse effect of air pollution. These type of pollutants include Sulphur Dioxide (SO_2), Nitrogen Dioxide (NO_2), Carbon Monoxide (CO), Hydrocarbon and Particulate Matter (PM). While secondary pollutants are pollutants manufactured in the air and are responsible for most of the smog, haze and eye irritation and for many of the form of plant and material damage attributed to air pollution. Ozone (O_3) is secondary pollutant, which is formed in the air as a result of chemical reactions.

Although air quality in Malaysia as a whole is not yet critical, the situation in large urban centers has already posed a threat to health and quality of life. To evaluate the exposition to atmospheric pollution, an accurate knowledge of the spatial distribution of pollutants in urban area is required. Manual technique of monitoring air quality involves the measurement of pollutant concentrations within the area of measuring station. A more practical approach is needed to study air quality at micro-scale. One possibility is the use of remote sensing technology. As stated by Jensen (2000) remote sensing is also different from the other mapping sciences such as cartography or GIS because they rely on data produced elsewhere. Remote sensing science yields fundamental scientific information.

Remote sensing represents a powerful technology for providing input data for measurement, mapping, monitoring and modelling within GIS context (Estes et al., 1984). There are many ways of integrating the techniques of satellite remote sensing and GIS for air quality monitoring. Ashalantha and Raghavswamy (2000) use remotely sensed data and GIS to generate air quality hot spot map. The image of IRS (PAN+LISS-III) was visually interpreted based on image characteristics such as tone, texture, shape, size, shadow and pattern as well as associated elements like location and association. These elements of interpretation helped in identifying and delineating thematic information related to urban land use or land cover (Raghavswamy *et al.,* 1998). While GIS techniques were applied to generate a 1 km buffer around the sampling location. The interpolation methods have been used to map the pollutant concentrations. Kanagroglou *et al*. (2002) used geostatistical interpolation method, kriging to improve the satellite derived of pollution maps. Universal kriging method has been used to produce reasonable estimate for missing value of patches of Aerosol Optical Thickness Value (AOTV), which correlate highly with air quality.

Earlier studies have revealed the possible relationship between the air pollution and satellite images (Sifakis *et al*., 1992; 1998; Retalis *et al*., 1999, Wald and Balleynaud, 1999; Ahmad and Mazlan, 1997; 2000; Ung *et al*., 2001, Kanaroglou *et al*., 2002). As mentioned in Sifalis *et al.,* (1992), high spatial resolution satellite data provide the opportunity to produce map indicating the horizontal distribution of the total aerosol loading over polluted areas. Ahmad *et al.*, (2000) utilized NOAA AVHRR data to determine haze API from forest fire emission during the 1997 thick haze episode.

Methodology

The methodology adopted for this study are divided into four main stages. The steps are as follows: i) data acquisition, ii) preliminary data processing, iii) data analysis and iv) generation of air quality maps. The primary data for this research is the air quality data and the Landsat ETM+ satellite image. Air quality data of eight (8) Continuous Air Quality Monitoring (CAQM) stations located at the following locations i.e. Kuala Lumpur, Gombak, Kajang, Klang, Petaling Jaya and Shah Alam are obtained from Alam Sekitar Sdn Bhd (ASMA). Although the data for the five (5) components of pollutants i.e. Sulphur Dioxide [SO_2], Nitrogen Dioxide [NO_2], Ozone, Carbon Monoxide [CO], Particulate Matter (particle less than 10 micron in size) [PM10] as concern in Malaysia Air Pollution Index (API) are obtained only four (4) are of main concern in this study

For this study, full scene Landsat 7 ETM+ image acquired on the 15th of July 2000, which covers Selangor and part of Negeri Sembilan is used. The ETM+ sensor consists of eight (8) bands (visible bands :- Band 1: 0.45 – 0. 52 μm, Band 2: 0.53 – 0.60μm and Band 3: 0.63 – 0.69μm; near-infrared bands :- Band 4: 0.76 – 0.90 μm; short-wave infrared band – Band 5: 1.55 – 1.75μm and Band 7: 2.09 – 2.35μm, thermal band – Band 6: 10.4 – 12.5μm and a panchromatic band. The spatial resolution for thermal band (Band 6), multispectral bands and panchromatic band are 60, 30 and 15 metres respectively. The thermal Infra Red (IR) band is particularly thermal mapping application. These data are particularly important in determining the relationship between the image Digital Number (DN) and air pollution index reading.

Full scene image is later subset into two different images, i) Klang Valley Region and ii) Shah Alam area. The image of the Klang Valley Region covers approximately 60 km from east to west and 40 km from south to north. The Klang Valley Region consist of Federal Territory (Kuala Lumpur), Klang, Petaling Jaya, Gombak, and Hulu Langat. Landsat 7 ETM+ colour composites images (Band 4, Band 3, Band 2) and (Band 3, Band 4, Band 5) of the Klang Valley Region are shown in Figures 1 and 2. The subset image of Shah Alam area is shown Figure 2. The administrative boundary layer of the Klang Valley Region and Shah Alam together with the location of CAQM stations (refer to Table 1) are superimposed onto the image (Figure 2). The thermal band of the study areas are shown in Figure 3. The air quality data at 10 am on the date the satellite data is acquired is given in Table 2. Figure 4 shows the locations of CAQM stations within the study area.

Table 1: Coordinates of CAQM Stations

Stn ID	CAQM Station	Eastings (m)	Northings (m)
1	Kuala Lumpur	347259.642	412068.035
2	Gombak	360939.767	406228.961
3	Kajang	331326.620	416100.169
4	Klang	333215.041	379076.943
5	Petaling Jaya	344174.824	412060.761
6	Shah Alam	340586.252	390560.750
7	Nilai	312145.011	424221.805
8	Seremban	301369.706	441298.924

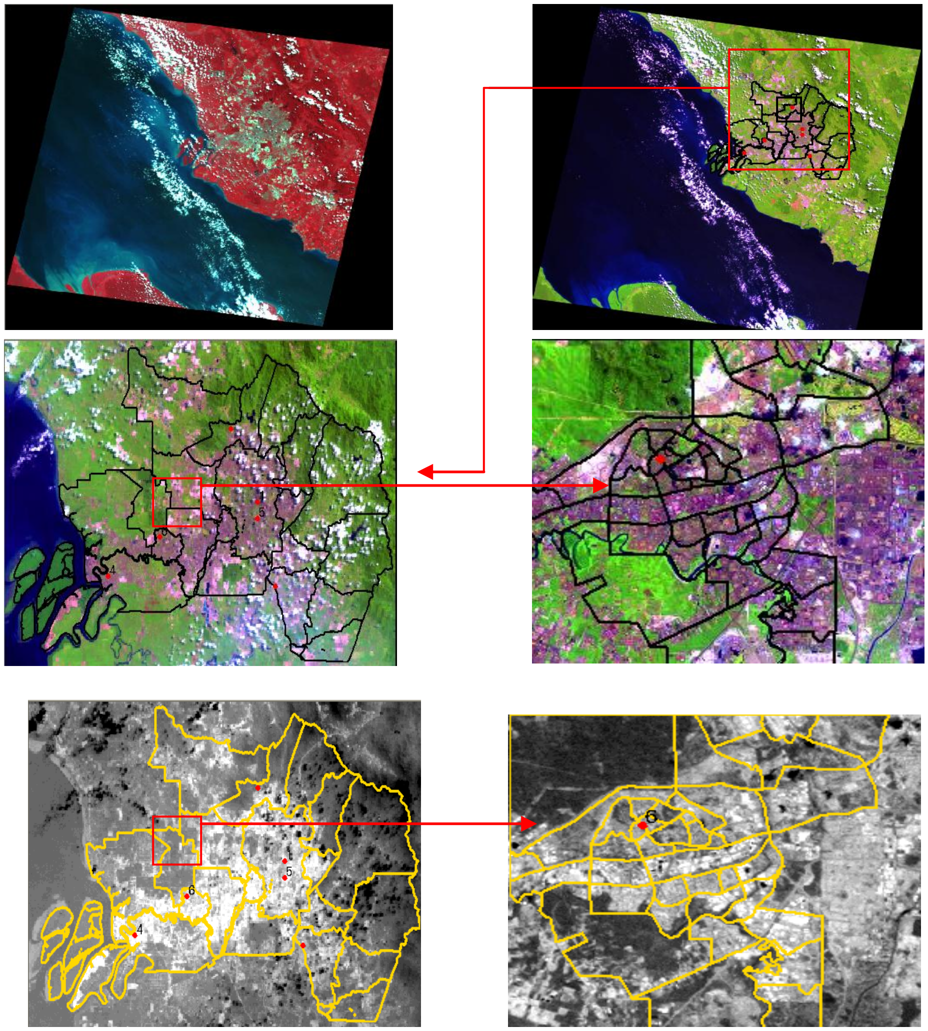


Figure 3: Sub-scene image (thermal band) of the Klang Valley Region and Shah Alam area

Table 2: Air quality reading at various CAQM Stations

CAQM Station	K.L	Gombak	Kajang	Klang	P.Jaya	S.Alam	Nilai	Seremban
Pollutant/Time	10:00	10:00	10:00	10:00	10:00	10:00	10:00	10:00
SO_2 ($\mu g/m^3$)	48.462	29.615	16.154	37.692	13.462	61.923	35.000	5.385
NO_2 ($\mu g/m^3$)	114.824	86.588	33.882	47.059	48.941	42.824	35.765	20.706
CO ($\mu g/m^3$)	5.052	2.415	1.3417	3.908	N/A	2.228	1.027	0.77
PM_{10} ($\mu g/m^3$)	112	116	85	154	93	108	123.000	65.000

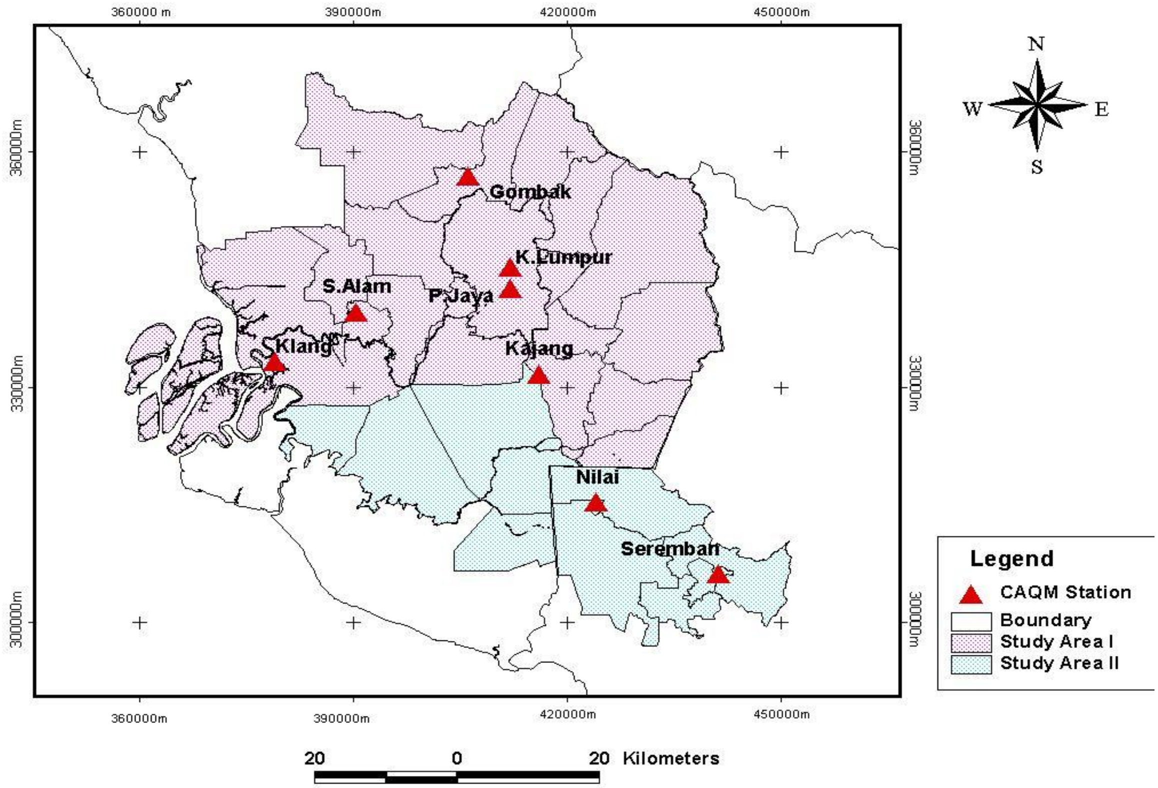


Figure 4: Location of CAQM stations within study area

Preliminary data processing involved digital image analysis to extract the DN of thermal and visible (1, 2, 3) bands at the eight CAQM stations. Once the DN of thermal band have been extracted, the DNs at the CAQM stations are compared to the air quality readings of their corresponding stations. Regression analysis is carried out to determine the relationship between DN and different air quality parameters. In order to generate initial virtual stations, the DN of bands 1, 2 and 3 at the CAQM stations are extracted and new locations within the image with similar DN are identified. To positively identify virtual stations, ground verification on the initial virtual stations is carried out. The main criteria for selecting these stations are i) point with the same spectral reflectance or DN of the CAQM stations, ii) point with the same surrounding features (land use) as the CAQM stations and iii) a minimum of four (4) stations within an area of 2.5 km2 (Sifakis, 1992). From these virtual stations, the air quality maps of various pollutant parameters are derived.

Results and Discussion

Linear regression analysis has been carried between DN of satellite imagery and four different types of pollutant parameters, that are SO_2, NO_2, CO, and PM_{10}. Different CAQM station configurations were used. Table 3 shows the result of linear regression analysis between four (4) types of air pollutants parameters at 10am and DN of the Landsat thermal IR band of 15th July 2000. The strong correlation are obtained when four (4), five (5) and six (6) CAQM station configurations are used. The strongest correlation (based on the configuration of (6) CAQM stations) is between DN and SO_2, followed by DN and CO, DN and NO_2, and DN and $PM_{10.}$ The correlation between DN and SO_2 range from $R^2 = 0.999$ and $R^2 = 0.765$. The weakest correlation

$R^2 = 0.765$ is when 7 stations are used. For correlation between DN and CO the range is between $R^2 = 1.000$ and $R^2 = 0.689$. The strongest correlation is obtained when only four (4) CAQM stations are used. The correlation between DN and NO_2 range from $R^2 = 0.984$ and $R^2 = 0.341$. For correlation between DN and PM_{10}, the correlation range between $R^2 = 0.991$ and $R^2 = 0.425$. The correlation between DN and O_3 range from $R^2 = 0.993$ and $R^2 = 0.324$. When all the stations are used, the correlation is reduced ($R^2 = 0.341$ for NO_2 and $R^2 = 0.425$ for PM_{10}).

Table 3: Regression analysis between pollutants and DN for dataset dated 15th July 2000

Pollutants	**Regression**	**R^2**	**Station ID**
SO_2	y = 0.1271x + 141.18	0.999	2,3,4,8
	y = 0.1322x + 141.12	0.987	2,3,4,7,8
	y = 0.1724x + 140.38	**0.916**	1,2,3,4,7,8
	y = 0.1057x + 141.89	0.607	1,2,3,4,6,7,8
	y = 0.1294x + 142.12	0.765	1,2,3,4,5,6,7,8
NO_2	y = 0.054x + 143.76	0.984	1,6,7,4
	y = 0.1983x + 137.33	0.883	3,4,5,6,8
	y = 0.0757x + 141.78	**0.810**	1,3,4,6,7,8
	y = 0.0756x + 142.14	0.711	1,3,4,5,6,7,8
	y = 0.0525x + 143.3	0.341	1,2,3,4,5,6,7,8
CO	y = 1.8748x + 140.51	1.000	1,2,3,8
	y = 0.1322x + 140.81	0.964	1,2,3,6,8
	y = 0.1724x + 141.01	**0.888**	1,2,3,4,6,8
	y = 0.1057x + 142.2	0.689	1,2,3,4,6,7,8
PM_{10}	y = 0.1271x + 137.53	0.991	2,4,7,8
	y = 0.1322x + 139	0.896	2,3,4,7,8
	y = 0.1724x + 139.29	**0.766**	2,3,4,6,7,8
	y = 0.1057x + 137.52	0.366	1,2,3,5,6,7,8
	y = 0.1294x + 138.83	0.425	1,2,3,4,5,6,7,8

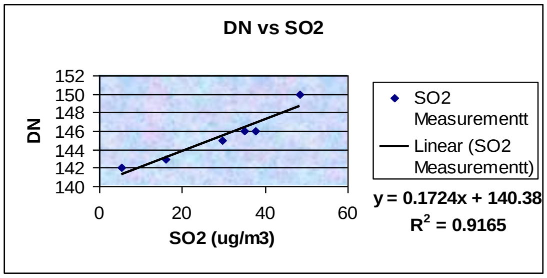
Based the regression analysis carried out, four regression models are selected. All these models are derived from six (6) CAQM station configuration. The selected models are used to derive the pollutant values (i.e. SO_2, NO_2, CO, and PM_{10}) for virtual air quality station are as follows:

For SO_2, $y = 0.1724x + 140.38$
For NO_2 $y = 0.0757x + 141.78$
For CO $y = 0.1724x + 141.01$
For PM_{10} $y = 0.1724x + 139.29$

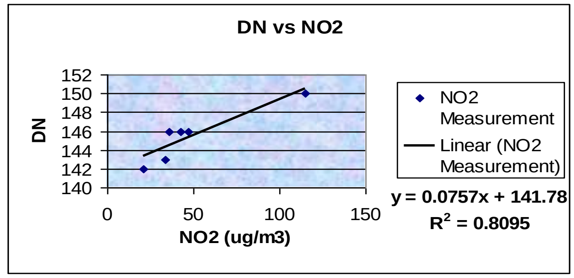
where, y - DN of CAQM station
x - Pollutant value

The regression analysis is illustrated graphically by using trendlines in scatter plot to display trends in data and to analyze the relationship between DN and air quality data. Figure 5 graphically shows the regression analysis between DN and pollutant concentration based on configuration of Kuala Lumpur, Gombak, Kajang, Klang, Nilai and Seremban (i.e. configuration of 1, 2, 3, 4, 7, 8) CAQM stations. A high correlation is found between DN and SO_2 ($R^2 = 0.916$) (refer to Figure 5 a). Figure 5 b) shows a linear regression analysis between DN and NO_2 from the CAQM stations of Kuala Lumpur, Kajang, Klang, Shah Alam, Nilai and Seremban (i.e.

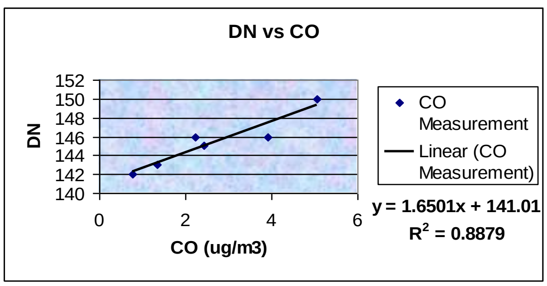
configuration of 1, 3, 4, 6, 7, and 8). The correlation coefficient R^2 is 0.810. Another regression has been carried between DN of thermal IR band and CO (refer to Figure 5 c)). The value of R^2 is 0.888. The result obtained is based on CAQM stations of Kuala Lumpur, Gombak, Kajang, Klang, Shah Alam and Seremban. (i.e. configuration of CAQM stations 1, 2, 3, 4, 6, 8). The correlation between DN and PM_{10} (R^2 = 0.766) is shown in Figure 5 d). The stations involved in the regression analysis are Gombak, Kajang, Klang, Shah Alam, Nilai and Seremban (i.e. 2, 3, 4, 6, 7 and 8).



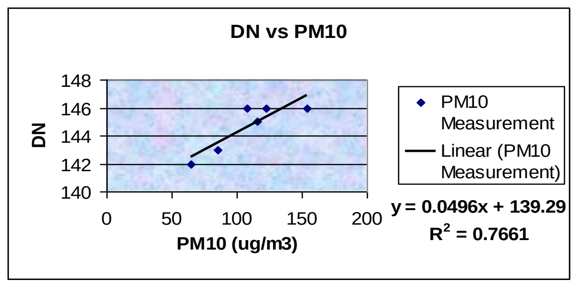
a)



b)



c)



d)

Figure 5: Relationship between DN and a) SO_2, b) NO_2, c) CO and d) PM_{10}

Maps of pollution concentration of SO_2, NO_2, CO and PM_{10} generated using the kriging interpolation method based on six CAQM stations located within the Klang Valley Region are shown in figures 6 a), 6 b), 6 c) and 6 d) respectively. The introduction "virtual stations" (57 stations) produce a more realistic distribution of air pollution within the Shah Alam area. Figures 7 a), 7 b), 7 c) and 7 d) show the pollution concentration maps of SO_2, NO_2, CO and PM_{10} generated based on 57 virtual stations respectively. The concentration of SO_2, NO_2, CO and PM_{10} are highest in industrial zones (especially Sections 15, 17 and Batu Tiga). The lowest concentration is mainly found within the forested areas.

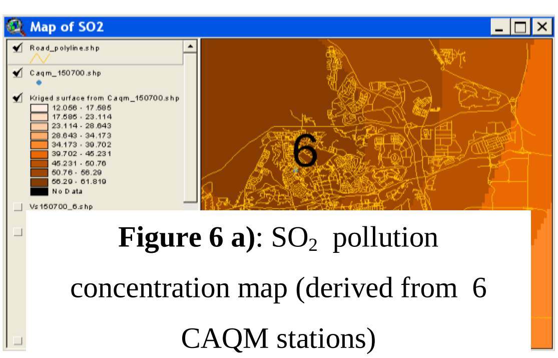


Figure 6 a): SO_2 pollution concentration map (derived from 6 CAQM stations)

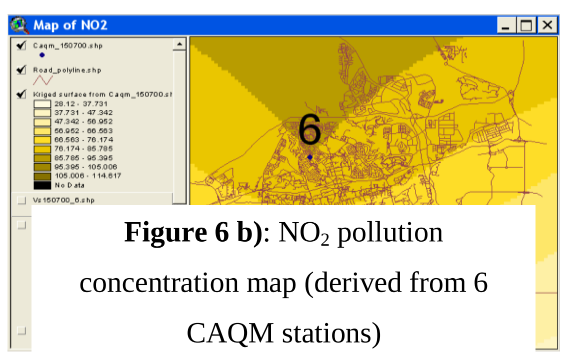


Figure 6 b): NO_2 pollution concentration map (derived from 6 CAQM stations)

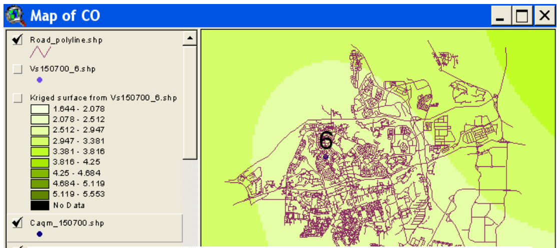


Figure 6 c): CO pollution concentration map (derived from 6 CAQM stations)

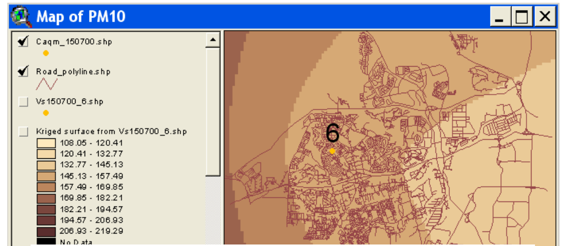


Figure 6 d): PM_{10} pollution concentration map (derived from 6 CAQM stations)

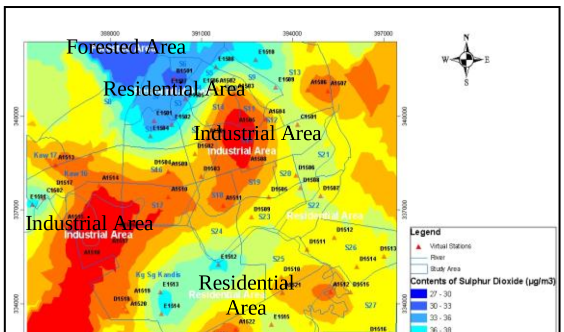
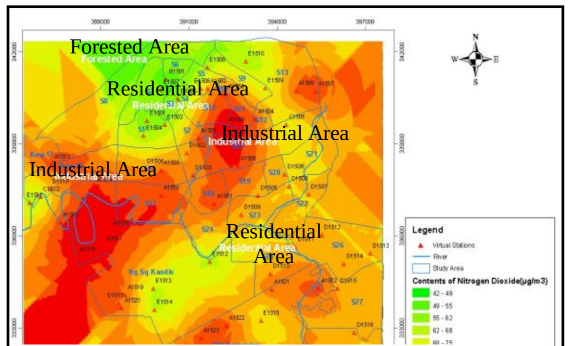


Figure 7 a): SO_2 pollution concentration map (derived from 57 virtual stations)



- **Figure 7 b)**: NO_2 pollution concentration map (derived from 57 virtual stations)

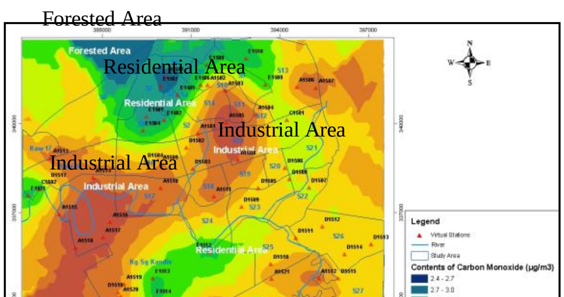


Figure 7 c): SO_2 pollution concentration map (derived from 57 virtual stations)

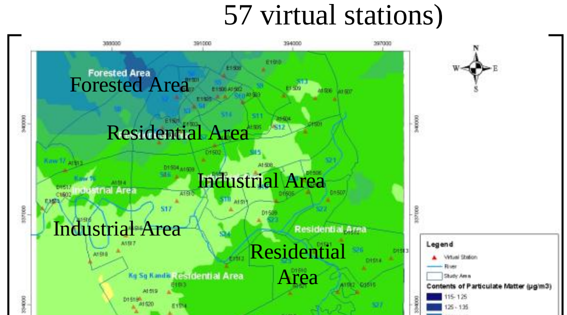


Figure 7 d): PM_{10} pollution concentration map (derived from 57 virtual stations)

Conclusions

Although this is an ongoing research, initial findings from this research can be summarised as follows:-

- there is a strong correlation between satellite image DN and the pollutants

- satellite-based image provide a low-cost means of producing air quality map of an area (especially at micro-scale level)
- the introduction of "virtual stations" gives a more realistic distribution pattern of the pollutants

Results from this research have shown that the satellite image especially LANDSAT ETM+ can be used to produce detailed air pollution map. Further study should include independent ground checks to verify the validity of the proposed method. Satellite images from other sensors such as Quickbird, IKONOS and SPOT 5 can also be tested.

Acknowledgements

Special thanks are due to the Malaysian Centre of Remote Sensing (MACRES) for providing the LANDSAT images and to Alam Sekitar Sdn. Bhd. (ASMA) for the providing air quality data of CAQM stations.

References

Afroz. R., Mohd Nasir Hassan, Noor Akma Ibrahim, 2003. Review of Air Pollution and Health Impacts in Malaysia. *Environmental Research*. Vol 92, Issue2. pp. 71-77.

Ahmad A. and M. Hashim, 2000. Determination of Haze API from forest fire emission during the 1997 thick haze episode in Malaysia using NOAA AHRR data. *Malaysian Journal of Remote Sensing & GIS*, 1, pp. 77-84

DEPARTMENT OF ENVIRONMENT MALAYSIA, 1983. *Report of Air Quality Management Study for Klang Valley Region*

Kanaroglou P.S., N.A. Soulakellis, Sifakis N.I., 2002. Improvement of satellite derived pollution maps with the use of a geostatistical interpolation method. *Journal of Geographical Systems*, (4), pp. 193-208.

Retalis A., Cartalis C. and Athanassiou E., 1999. Assessment of the distribution of aerosols in the area of Athens with the use of Landsat Thematic Mapper data. *International Journal of Remote Sensing*, (20)5, 939-945

Sally L. J. and A.J Rossini, 1996. Use of Kriging Models To Predict 12-Hour Mean Ozone Concentrations In Metropolitan Toronto-A Pilot Study. *Environment International*. Vol 22. No . 6, pp. 677-692, 1996.

Sham S., 1987. *Urbanization and the Atmospheric Environment in the Low Tropics*. Universiti Kebangsaan Malaysia.

Stern A. C, Boubel R. W., Turner D.B, Fox D.L., 1984. *Fundamentals of Air Pollution*. Academic Press Inc

Tayanc M., 2000. An Assessment of Spatial and Temporal Variation of Sulphur Dioxide Levels Over Istanbul, Turkey. *Environmental Pollution*. Vol 107, Issue 1. pp. 61-69.

Ung A., L. Wald, T. Ranchin, C. Weber, J.Hirsch, G.Perron and J.Kleinpeter, 2001. Satellite Data for Air Pollution Mapping Over a City – Virtual Stations. *Proceedings of the 21 th EARSel Symposium, Observing our environment from space: new solutions for a new millennium, Paris, France*, pp. 147-151.

Wald L. and J. –M. Baleynaud, 1999. Observing air quality over the city of Nantes by means of LANDSAT thermal infrared data. *International Journal of Remote Sensing*, (20) 5, pp. 947-959.