DATA CAPTURE AND DATA MINING OF URBAN AIR POLLUTION: THE BUILDING-BASED APPROACH

U. W. Tang
Department of Civil and Environmental Engineering, University of Macau, Macau (PR China).
Email: uwtang@umac.mo; uwtang@gmail.com

ABSTRACT

The method and accuracy of data capture dominate the spatial distribution of urban air pollution. Due to limited budget, installation space, and labor resource, permanent or temporary air pollution monitoring sites are very scattered. Air quality assessment of a city based on scattered monitoring sites may be incorrect because non-homogeneous distribution of air quality is neglected. Therefore, a number of model systems have been developed to estimate urban air quality at unsampled sites. In this paper, representative air quality model systems, their data captures and their applications are reviewed, which show that the input/output spatial data are commonly stored in regular grids with resolutions of 1-2 km, regardless of the complexity of urban form. Recently, a model system which can estimate air quality (and noise) in front of individual buildings along both sides of the road is developed. Compared with the grid-based approach with spatial resolutions of 1-2 km, the present building-based approach can predict the complex spatial variation of traffic emission, urban geometry, dispersion and air (noise) pollution. The results show that the building-based approach may open an innovated methodology in data mining of urban spatial data for environmental assessment.

1. INTRODUCTION

The method and accuracy of data capture dominate the data output (i.e., spatial distribution) of urban air pollution. Due to limited budget, installation space, and labor resource, permanent or temporary air or noise pollution monitoring monitoring sites are very scattered. Air quality assessment of a city based on scattered monitoring sites may be incorrect because non-homogeneous distribution of air quality is neglected. Therefore, a number of model systems have been developed to estimate urban air quality or noise at unsampled sites.

Air quality modeling is performed with computer programs that solve the mathematical equations and algorithms which describe the formation, transport and fate of air pollutants. The necessary input data includes emission parameters, meteorological conditions, and topography. As air quality modeling of a large urban area requires processing a large amount of complex geographically referenced data, it has a strong desire to incorporate air quality or noise models with a Geographic Information System (GIS) that enables capture, modeling, manipulation, retrieval, analysis, and presentation of geographically referenced data. Today, GIS has been extensively used for environmental modeling purposes (National Center for Geographic Information and Analysis, 1996).

Nevertheless, the scale or spatial resolution is still an important research topic in disciplines which underlying researches involving geographic information. Researches in environmental and urban management have shown that the outcomes can be altered according to the selection of scale in the studies (Gontier M., 2007; Partidário, 2007; João, 2002). The importance of scale issues has led some researchers to propose a new science of scale, which would include the studies of measures or properties which are invariant with respect to scale, the measurements for the impact of scale alternation, the influences of scale as a parameter in process models, and the implementation of multi-scale approaches (Goodchild and Quattrochi, 1997; Montello and Golledge, 1999).

This paper focuses on the use of an innovated approach, namely “the building-based approach”, to increase the spatial resolutions of input/output data for the modeling of urban air pollution, so that the data mining in urban and air pollution studies can be improved. In Section 2, representative air quality model systems, their data captures and applications will be reviewed, which show that the input/output spatial data in GIS are commonly stored in regular grids with resolutions of 1-2 km, without considering the complexity of urban form and air pollution distribution inside each grid. In Section 3, an innovated building-based approach for air quality modeling will be introduced. The data capture of urban geometry by the building-based approach will be described in details in Section 4. Applying the building-based approach, case studies in air quality modeling in Macau will be introduced in Section 5.

2. DATA CAPTURES & OUTPUTS OF REPRESENTATIVE MODEL SYSTEMS: THE GRID-BASED APPROACH

A number of air quality model systems such as the Atmospheric Dispersion Modelling System (ADMS-urban) (McHugh et al., 1997), the Air Quality Information System (AirQUIS) (Bøhler et al., 1998), the Environmental Management (EnviMan) (Kadikis et al., 2002), and the Urban Simulation Model (Urban SIM) (Alberti and Waddell, 2000) have been developed to predict the spatial distribution of air quality of a city. One of the most representative air quality model systems is the desktop ADMS-Urban developed by the Cambridge Environmental Research Consultants and the U.K. Meteorological Office (McHug et al., 1997). ADMS-Urban is based on ADMS that was originally developed to assess point source pollution using a Gaussian dispersion model. For air quality management applications in an urban environment, ADMS has been developed as a multi-source (i.e., point, line and area sources) dispersion model system ADMS-urban which links to an emissions inventory database to provide emissions from all source sectors and integrates a Gaussian dispersion model and a street canyon model (OSPM) (Hertel and Berkowicz, 1989). Particularly, traffic sources are entered in terms of number, type, and speed of vehicles.

Blair et al. (2003) have linked ADMS-urban with ArcView GIS to predict the spatial distributions of concentrations of nitrogen dioxide (NO_2) and particulate matter (PM_{10}) in London, which provides a typical example of data capture and data output of ADMS-urban. The point, line and area emission sources in the Greater London emissions inventory database were captured by ArcView GIS and then input to the dispersion model as the aggregated total on a 1×1 km grid square basis. The urban geometric characteristic of London, i.e., one of the key components of the modeling in terms of the effects on wind speed and turbulence, was represented by a universal and highly simplified parameter (i.e., surface roughness = 1 m). For data output, Blair et al. (2003) mentioned that it was not possible to model the whole of the Greater London area which covers 59 km x 45 km in one calculation at the highest spatial resolution of ADMS Urban. In order to achieve suitable resolution on the output maps, the modeled area was split up into 23 sections of varying size in which the central London areas were modeled with higher output resolution (e.g. 1 km x 1 km) than the outer areas. The work of Blair et al. (2003) is a typical example showing the very rough and homogeneous data capture of urban geometry for modeling of urban air quality distribution.

Urban geometry and urban land use are two basic characteristics of urban form. Considering the importance of urban form on sustainable development recognized in recent years (National Physical Planning Agency, 1991; Department of the Environment, 1992), a few model systems have been built to study air quality distribution in different urban forms (Newton, 1997; Marquez and Smith, 1999; Borrego et al., 2006). These studies focused on the influence of urban land use on the intensities and distributions of traffic, domestic and industrial emissions and their environmental consequences, i.e., distributions of air pollution. However, urban geometry or street configuration (another key characteristic of urban form) is neglected because the distributions of air pollution were estimated by dispersion models with grid resolutions of kilometers. Details will be given as follows.

The modeling framework developed by Newton (1997) and Marquez and Smith (1999) integrates land use, transport and airshed models It consists of five components. A land use/transport component specifies land use scenarios and simulates traffic conditions in each urban form. An emissions component calculates distribution of emissions of point, line, area and biogenic sources. A meteorological component receives mesoscale meteorological data and generates climatic conditions over the study area. An airshed model uses the forecast meteorological conditions from the meteorological component and the gridded, hourly-averaged forecast emissions from the emissions component to track the movement of contaminants and calculate the rate of chemical reaction that leads to the formation of photochemical smog and secondary aerosols. Particularly, a GIS/database component is built for capturing data required by the other four components.

Based on the framework, Newton (1997) and Marquez and Smith (1999) forecasted the air quality in Melbourne, Australia for 2011 by considering four urban form scenarios. The spatial resolution of data capture and data output was on a regional scale of 2 km x 2 km. In the first urban form, the population growth and urban land use were concentrated in inner city.

This led to the assumptions of heavy traffic congestion, significantly high pollutant emissions, and thus high air pollutants. In the rest three urban forms, namely compact, edge and corridor urban forms, the population growth and urban land use were concentrated in the inner city and suburbs, in subcentres around the metropolitan area, and along three transportation corridors connected to the central business district (CBD), respectively. The edge and corridor urban forms were supported by the radial rail transportation which delivered environmental and efficiency benefits. Modeling results showed that the compact, edge, and corridor urban forms led to less traffic congestion, lower pollutant emissions, and thus better urban air quality when compared to the business-as-usual urban form in Melbourne.

Borrego et al. (2006) applied a photochemical model system MEMO/MARS (ITT, 1994) to simulate the potential impacts of three imaginary urban forms, namely disperse, corridor, and compact urban forms, on urban air quality. The system includes a meteorological model MEMO and a photochemical model MARS. The MEMO model is a non-hydrostatic prognostic model developed to simulate the wind flow over complex terrain (ITT, 1994). The MARS model describes the dispersion and chemical transformation of air pollutants in a three dimensional region (Moussiopoulos, 1995). In the study of Borrego et al. (2006), traffic emission data required by the MARS model were estimated according to the total number of vehicles, total kilometers travelled, and vehicle emission factors. All emission data were considered as area sources and were represented by a regular matrix.

The disperse urban form imagined by Borrego et al. (2006) by was characterized by low density and large area requirements. The urban area was separated into distinct zones for residential, commercial or industrial uses, with the consequent high vehicle dependence but the lowest emissions per area. In the corridor urban form, urban growth was located in linear corridors with origin in the city centre. This was supported by high quality transport infrastructure (highways) and characterized with the highest emission rates. The compact urban form was characterized by high density, less area requirements, and mixed land use. Complementary functions (housing, shopping and offices) were located close together, which allowed the reduction of travel length and number of trips, and thus led to lower emission rates per inhabitant. Based on the predicted air quality distributions with a spatial resolution of 2 km x 2 km, Borrego et al. (2006) concluded that the compact urban form with mixed land use provided better air quality when compared to the disperse urban form with segregated land use and the corridor urban form equipped with intensive transport structures.

To this end, it is shown that traditional air quality model systems are designed for applications in regional scale with input/output resolutions in kilometers, regardless of the complexity of urban geometry within the urban areas. Even in the studies of the impact of urban form on air quality distribution, their objectives only focused on the impact of various land use patterns on emissions and air quality distribution. The impact of urban geometry on air pollution dispersion was considered very roughly.

3. THE BUILDING-BASED APPROACH

Air quality monitoring studies co-conducted by the author in Macau (the Macau Peninsula), an urban area with highly compact urban forms, have shown the significant spatial variation of air pollution within kilometer squares. In an area of 1 km × 4 km, the roadside polycyclic aromatic hydrocarbons levels of dustfall samples at eleven sites varied from 2.72 to 24.83 µg/g (Qi et al., 2001a,b). Even in an neighborhood scale of 670 m × 200 m, toxic volatile organic compounds BTEX at 18 sampling sites varied to a great extent, i.e., the highest total BTEX was 136 times higher than that of the lowest (Chen et al., 2000). It is obvious that regional scale model systems with input/output resolution in kilometers cannot precisely predict the complex spatial variation of urban geometry, dispersion and air quality in Macau.

Therefore, a model system with higher spatial resolution is necessary. Some researchers have developed high temporal/spatial resolution systems for modeling air quality or noise (Jensen, 1998, 1999; Borrego et al., 2000, 2003, 2004; De Ridder et al., 2004). In particular, Jensen (1998, 1999) developed a prototype air quality model system based on GIS to support local authorities in air quality management for big Danish cities. The system integrated digital maps, administrative databases, an Operational Street Pollution Model (OSPM) (Hertel and Berkowicz, 1989), an urban landscape model and ArcView GIS for air quality and exposure estimation at the street address level. Compared with existing urban air quality management tools such as the Atmospheric Dispersion Modeling System (ADMS-urban) (McHugh et al., 1997), Jensen’s model system has higher spatial resolution, makes use of digital maps and administrative databases for automatic generation of street configuration data, adds standard GIS features, and provides improved exposure assessment (Jensen, 1999).

Building on the work of Jensen (1998, 1999), a prototype air quality model system has been developed for Macau (Tang and Wang, 2007). As shown in Fig. 1, the present model system integrates the operational air pollution model OSPM (Hertel and Berkowicz, 1989), the road traffic noise model CRTN (Department of Transport, Welsh Office, 1988), digital maps of the road network, building layout and topographic information, and an urban landscape model. The air quality in front of a building in Macau is determined by use of the OSPM model. The OSPM model is a practical street pollution model, developed by the Department of Atmospheric Environment, National Environmental Research Institute, Denmark (Hertel and Berkowicz, 1989). Concentrations of exhaust gases are calculated by a combination of a simple plume model for the direct contribution and a box model for the recirculating part of the pollutants in the street. In this study, the OSPM model programmed in FORTRAN is compiled into an executable module for integration with ArcView.

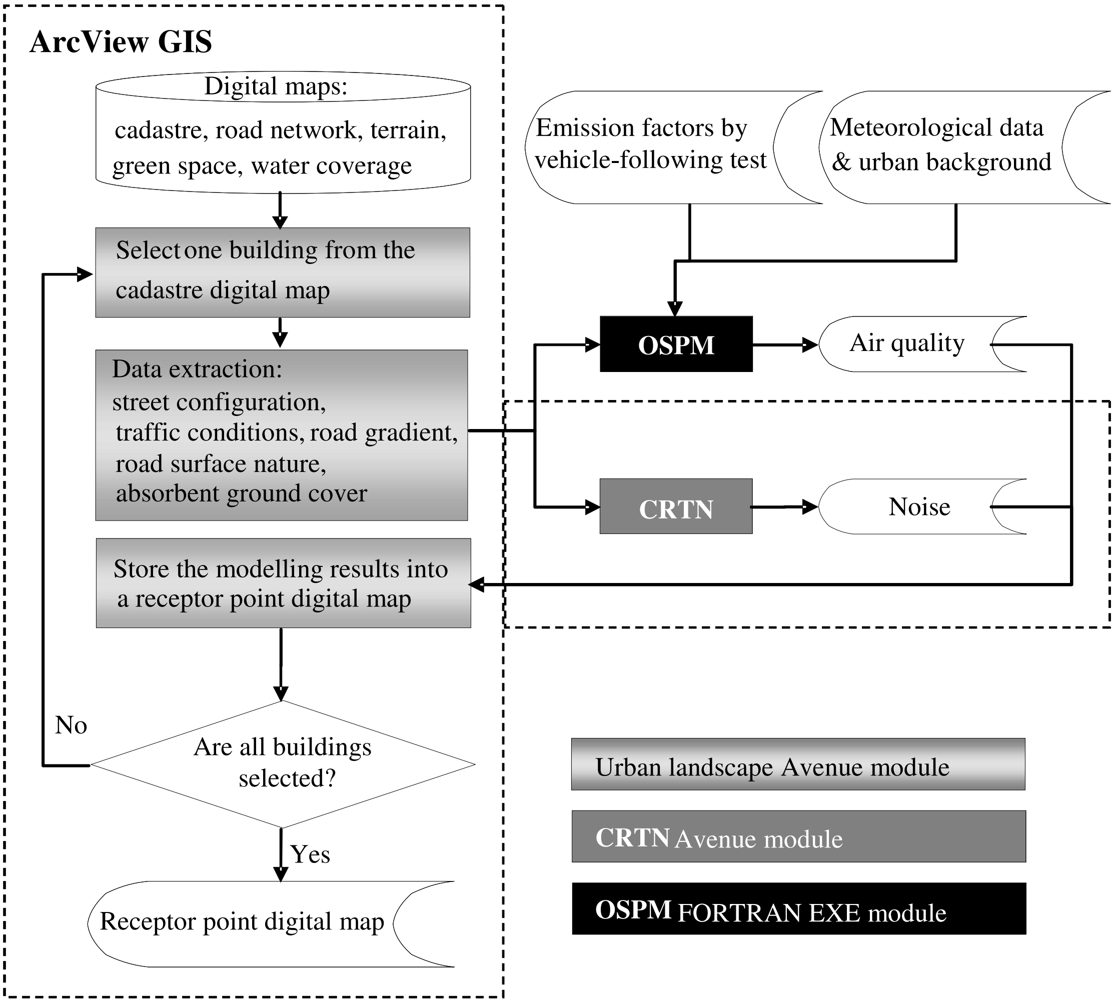


Fig. 1. Structure of air quality and noise model system.

The noise level in front of a building in Macau can be calculated using the CRTN model available in the Department of Transport's Technical Memorandum in the United Kingdom (Department of Transport, Welsh Office, 1988). The method assumes typical traffic and noise propagation conditions that are consistent with moderately adverse wind velocities and directions during the specified periods. In this study, the CRTN method is programmed in Avenue which is a scripting language of ArcView. ArcView is a commercially available, desktop GIS. The required input parameters such as traffic conditions, road gradient, road surface nature, absorbent ground cover percentage and street configurations are extracted from digital maps in ArcView format, namely themes.

The OSPM and CRTN models are integrated with ArcView GIS by a prototype urban landscape model programmed in the Avenue language. Avenue provides the customization and development environment for integrating ArcView with other applications (see Section 4 for details).

In air quality modeling, vehicle emission factors are essential input parameters for the OSPM. To obtain the emission factors for each category of vehicle (motorcycle, passenger car, taxi, truck and bus), vehicle chase tests were conducted under traffic congestion conditions at the urban hot spots in Macau between 24 February 2004 and 25 April 2004 (Tang and Wang, 2006). In comparison with the traditional emission tests such as chassis dynamometer and idle-emission tests, vehicle chase tests reflect real-world situations and more vehicle samples are available because it is not necessary to obtain the cooperation of the vehicle owners.

The major differences between the present model system and that developed by Jensen (1998) are:

Traffic noise prediction: Jensen’s (1998) model system is only for prediction of air quality distribution while the present model system can also predict traffic noise distribution at the same receptor point. Evaluation of the traffic noise modeling in Macau has been given in Tang and Wang (2007) and will not be repeated here.

Vehicle emission factors: Jensen’s (1998) used the vehicle emission factor model (COPERT III) to obtain vehicle emission factors. In the present model system, vehicle emission factors are obtained from vehicle chase tests conducted in the urban hot spots in Macau. Results and evaluation tests have been presented in Tang and Wang (2006). Compared with the emission factors obtained by modeling methods such as the mobile source emission factor model MOBILE5 (Hao et al., 2003), Tang and Wang (2007) showed that those obtained from the present vehicle chase tests can better represent real-world situations in Macau.

Accuracy of data capture: Jensen (1999) greatly simplified the data capture of urban geometry (i.e., extraction of street configuration data), which is a key factor for the calculation of pollutant dispersion by the OSPM. As shown in Fig. 2a, the description of street configuration around a center point on the road is handled by 12 wind sectors, where each wind sector covers an angle of 30 degrees. The centerline of the wind sector is used to identify the building and the associated building height of that wind sector. The distance from the centre point on the road centre line to the building in each wind sector is also calculated. If the distance is more than 25 meters, it is assumed that the wind sector has no buildings.

This phenomenon has a number of limitations to describe the street configuration effectively and precisely. When a wind sector (such as wind sector 1 or 2) covers more than one buildings, only one building and its associated height can be identified by the centre line of the wind sector, regardless of the possible significant variation of building heights within the wind sector. In wind sector 3, there is an open space which will lead to modification of the calculation steps of the air pollution recirculation in the OSPM. Nevertheless, the centerline of the wind sector can only identify the building and the OSPM can only treat this wind sector with a homogeneous building height. Furthermore, wind sectors 4 to 6 involve redundant identification steps because they are associated with one building only. On the other hand, the present model system follows exactly the guideline of the OSPM manual to define the boundary of each wind sector by the exact location of each building or open space (see Fig. 2b), so that the variation of heights in each wind sector can be taken into account by the dispersion calculation of the OSPM (see the next section for details).

Fig. 2. Wind sectors generated from the urban landscape model by (a) Jensen (1998) and (b) the present study.

4. DATA CAPTURE OF THE PRESENT MODEL SYSTEM

As shown in Fig. 1, data capture of the present model system is conducted by the urban landscape model. When the urban landscape model is executed, a target building polygon is selected from the cadastral map in ArcView (see Fig. 2b). Based on the coordinates of the vertices of the target building polygon, the boundary segments of the building polygon are obtained. If the boundary segment faces a street aligning parallel to it, it is treated as a building façade which is exposed to direct traffic air and noise pollution and a target receptor point is created in the middle of the boundary segment. During this process, the proximity of a boundary segment and a street is determined based on their locations in the cadastre map and the road network map, which share a common coordinate system. The target receptor point is stored in a receptor point map. In this study, the heights of all the receptor points are assigned as 1.8 m to simulate the breathing and hearing position of a human.

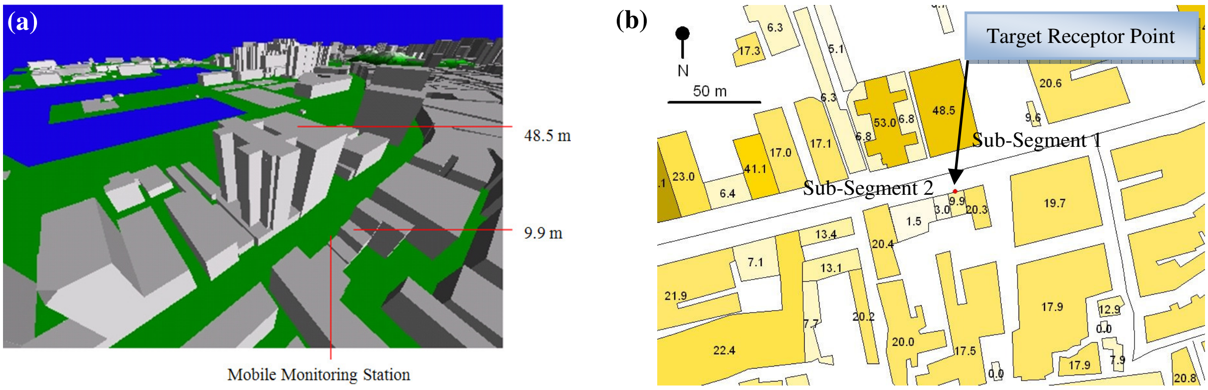
The spatial-related input data around the target receptor point are then extracted from the cadastre, road network, green space, and terrain maps for the OSPM and CRTN models. The spatial-related input data include the street configuration data, traffic data, road surface nature, road gradient, and absorbent ground cover. For a building polygon around the target receptor point, the coordinates of its vertices can be extracted from the cadastre map. Its position relative to the target receptor point can be described by a parameter of wind sector. As shown in Fig. 2b, the wind sector of a building polygon is defined with the origin at the road centre and the upper and lower bounds pointing from the origin to the two vertices of the building boundary segment.

The height of the building polygon is calculated by subtracting the elevation at the top of the building from that of the road surface in front of the building. The elevation at the top of the building is extracted from the attribute table of the cadastre map. The elevation of the road surface is extracted from the terrain map in ArcView. With the heights and the relative positions of the building polygons around the target receptor point, the distribution of the building heights around the receptor point can be obtained to determine the height of recirculation zone for the air quality modeling. The other 2D street configuration data required in the air quality and noise modeling, such as the length, width and orientation of the street in front of the target receptor point and the distance between the source line and the target receptor point can be determined by the urban landscape model based on the coordinates of the related features extracted from the cadastre map. The traffic data and the road surface nature on the street in front of the target receptor point can be extracted from the attribute records of the street in the attribute table of the road network map.

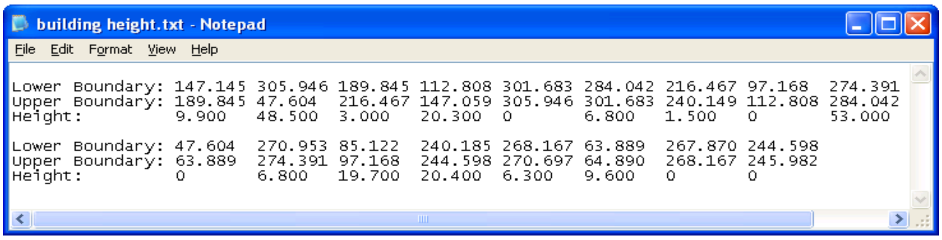
As shown in Fig. 1, the spatial-related data extracted from the digital maps in ArcView are input into the OSPM and CRTN models to simulate the air and noise pollution at the target receptor point. Since the OSPM model is integrated with ArcView as an executable module, the spatial-related input data extracted from the digital maps are stored in ASCII input files for the OSPM model. The non-spatial input data required by the OSPM model such as traffic emission factors, meteorological data and urban background concentrations, which have been prepared before the execution of the urban landscape model, are also stored in ASCII input files. On the other hand, the CRTN model is programmed in Avenue which is fully integrated into ArcView. Therefore, the required spatial-related input data extracted from the digital maps in ArcView can be input into the CRTN model as internal variables.

Without the need to generate and close spatial-related input files, the execution of the CRTN model is faster than that of the OSPM model. The modeling results of the OSPM and CRTN models, which include the gaseous concentrations of CO, NO, NO_2 and benzene, and the hourly noise levels of L_{10} and L_{eq}, are then stored in the attribute fields of the target point feature created in the receptor point map. Spatial-related input variables such as street configurations, traffic data, vehicular emissions, road gradient, road surface nature, and absorbent ground cover are also stored in the attribute fields of the receptor point. When all the buildings in the cadastral map are selected and manipulated, the urban landscape model stops the execution. All the modeling results in front of the building façades at roadside can then be accessed from the receptor point map in ArcView. The statistical/spatial relationships of the input values and output results can also be investigated.

To evaluate the air quality modeling of the present model system, the present model system is applied to predict air quality in a typical urban trunk road Rua da Ribeira do Patane in Macau and results are compared with measured values (Tang and Wang, 2007). A mobile monitoring station has been set up to measure continuous CO, NO_x (NO, NO_2), and O_3 at roadside. The immediate neighborhood is given in Fig. 3a, which shows that the measurement site is located in a street segment with complex building structures on both sides. To automatically generate street configuration input parameters required by the air quality modeling, a target building polygon which is located nearest to the mobile monitoring station is selected manually from the cadastral map in ArcView (see Fig. 3b). The urban landscape model of the present model system is then executed. The model selects the boundary segment of the target building polygon which faces the road and aligns parallel to it and creates a target receptor point in the middle of the boundary segment to simulate the position of the mobile monitoring station. The target receptor point is stored in a receptor point map.



(c)



```
Lower Boundary: 147.145 305.946 189.845 112.808 301.683 284.042 216.467 97.168  274.391
Upper Boundary: 189.845 47.604  216.467 147.059 305.946 301.683 240.149 112.808 284.042
Height:         9.900   48.500  3.000   20.300  0       6.800   1.500   0       53.000

Lower Boundary: 47.604  270.953 85.122  240.185 268.167 63.889  267.870 244.598
Upper Boundary: 63.889  274.391 97.168  244.598 270.697 64.890  268.167 245.982
Height:         0       6.800   19.700  20.400  6.300   9.600   0       0
```

Fig. 3. The experimental set-up in Rua da Ribeira do Patane; (b) street configurations around the receptor; (c) heights (m) of buildings around the receptor and their locations extracted as wind sectors (degrees) with origin in front of the receptor.

The spatial-related input data around the target receptor point are then extracted automatically from the cadastre, road network and terrain maps. The road segment in front of the target receptor point is divided into sub-segment 1 and sub-segment 2 according to the location of the target receptor point (Fig. 3b). The orientation of the sub-segment 1 is calculated to be 76.7° in relation to north and it is assigned as the orientation of the street segment. The lengths of the sub-segment 1 and sub-segment 2 are found to be greater than maximum input value of 75 m and thus treated as 75 m. The building polygons on both side of the sub-segment 1 and sub-segment 2 are then selected automatically. Their positions relative to the target receptor point are extracted as the parameters of wind sectors and stored in an ASCII input file, see Fig. 3c. The heights of the associated buildings are then extracted as the elevation difference between the top of the building in the cadastre map and the road surface in terrain map and stored in the same text file. By manual inspection, these extracted spatial-related input data are all correct. The data extraction by the urban landscape model has taken less than a second, which is much faster than the manual inspection.

The extracted street configurations and traffic conditions are input into the OSPM model to simulate the air pollution in front of the target building. It is shown that the modeled CO and NO_x concentrations agree well with the measured values. The correlation coefficient of the measured and modeled CO (NO_x) concentrations is 0.84 (0.67) and the fractional bias is 0.20 (0.10). A total of 100% (95%) of the CO (NO_x) data points are located inside the 'factor of 2' lines. The statistical validation results show that the model performance is satisfactory with good agreement between measurements and modeled concentrations.

5. DATA MINING BY THE BUILDING-BASED APPROACH

Macau is the first place in China open to the West in the 16th century and thus it has just been awarded United Nations World Cultural Heritage status for the preservation of its dramatic mixing of eastern and western cultural relic sites. After the reunification of Macau from Portugal to China in December 1999, Macau has experienced a tremendous success of economic development. In 2007, Macau became the largest gambling hub in the world and the richest place in Asia (in terms of per capita gross domestic product). Regardless of the historic areas with historic relic sites or the modern areas with mega casino hotels, the urban forms in Macau are highly compact due to the high population density (38,000 inhabitants/km^2) and mixed-used development (i.e., a mixture of residential, commercial, industrial, or other land uses in a building or set of buildings).

Applying the developed model system, the spatial relationships among air quality, traffic noise, traffic flow and urban geometry in the historic or modern areas in Macau are investigated. In the modeling, Traffic Scenario 1 uses the real traffic volumes counted manually during 18:00–19:00 on working days in 2001. Modeled results reflect the combined influences of street geometries and traffic volumes on street environment. Traffic Scenario 2 applies the traffic volume and composition of a busy single-lane road in the historical area to all the roads. This scenario is intended to determine the influences of street configurations on street environment when the variation of traffic condition is neglected. Typical meteorological conditions are chosen, i.e., wind direction 355°, wind speed 2.25 m/s, temperature 19.4°C, and background CO 0.87 ppm (Tang and Wang, 2007).

5.1 High-Spatial Resolution of Air Quality Distribution

Recalling Section 2, data capture of urban geometry and traffic emission for modeling of air pollution distribution in a city are usually in a resolution of kilometer(s) (Blair et al., 2003; Newton, 1997; Marquez and Smith, 1999; and Borrego et al., 2006). Such a resolution is obviously not suitable for air pollution distribution in Macau (i.e., the Macau Peninsula), which has only 8.8 km^2. In the previous studies co-conducted by the author in Macau, a higher resolution of 300×300 m grids have been used in data capture of traffic emissions (see Fig. 4a) and the subsequent modeling of annual carbon monoxide (CO) distribution (see Fig. 4b) (Wu et al., 2002a,b). The modeling of CO distribution was accomplished by a steady-state, Gaussian plume model (ISCST3) (Wu et al., 2002b) for transportation distance up to 50 km. CO concentrations are modeled at 110 receptor points located in the centre of the 300×300 m grids. The height of the receptors was set to 1.5 m, which is approximately equivalent to the human respiration height. The emission and CO distributions in Figs. 4a,b show a directly perceived conclusion that higher air pollution exposures can be found in the areas with higher traffic emissions. However, the influence of urban geometry (i.e., street configuration) on CO dispersion inside a grid or between different grids were not taken into account because the characteristics of urban geometry was only described by a highly simplified parameter - surface roughness.

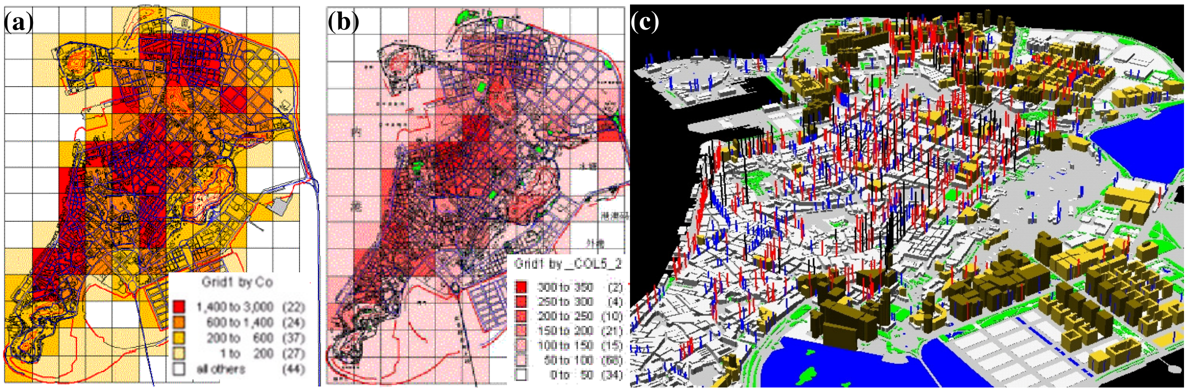


Fig. 4. (a) Distribution of CO emission by 300×300 m grids (Wu et al., 2002a), (b) CO concentrations by 300×300 m grids (Wu et al., 2002b), and (c) Modeled CO concentrations by building-based approach in Traffic Scenario 2 (Columns represent CO concentrations: blue 0–1.5 ppm, red 1.5–2.5 ppm, and black 2.5–5.6 ppm).

With the building-based approach invoked by the present model system, a total of 1343 receptor point entities are created automatically in front of individual buildings along both sides of the streets in ArcView GIS (see Fig. 4c). As the area of lands in Macau is only 7.7 km^2 (excluding the 1.1 km^2 of the reservoir and lakes), the average spatial resolution is 174 receptors per kilometer square, which is much higher than the previous studies with grid-based approach (Wu et al., 2002a,b; Blair et al., 2003; Newton, 1997; Marquez and Smith, 1999; and Borrego et al., 2006). In addition, as the street configuration and traffic emissions can be extracted building by building, the influence of the complex urban geometry and traffic conditions on the CO exposures at human respiration height can be taken into account. Human exposures in a city can then be investigated at the address level, which is more precise than those based on the roughly or unrealistic assumption of homogeneous distributions of urban geometry and air pollution inside the grids with resolutions of kilometers.

With the ArcView's 3D-Analyst extension, spatial relationships among air quality, noise, traffic flow and urban geometry can be presented in a 3D view by the author's model system (see Fig. 4c). Such 3D visualization is particularly useful to urban planners when they need to consider the influences of various urban form policies on street environment. Particularly, the highly compact urban forms in Macau may make the findings presented in the next section very relevant to them as they prepare for higher population densities in cities throughout the world.

5.2 Quantitative Analysis of Urban Form and Air Pollution

Recalling Section 2, the scenarios of urban forms were distinguished by land use patterns rather than urban geometry (Newton, 1997; Marquez and Smith, 1999; and Borrego et al., 2006). Similarly, in a noise pollution study co-conducted by the author in Macau, an attempt was made to link the measured urban noise with eight land use types, which were distributed in accordance with a grid resolution of 500 m × 500 m (see Fig. 5a, Bai et al., 2002). The noise in each land use type is assessed in terms of the value measured near the centre of each grid. Nevertheless, as the urban forms were distinguished by descriptive method (i.e., different urban land use types), these studies were restricted to descriptive analysis rather than quantitative analysis of the impact of urban forms. The urban geometry or street configuration, which is another key characteristic of urban form, was neglected.

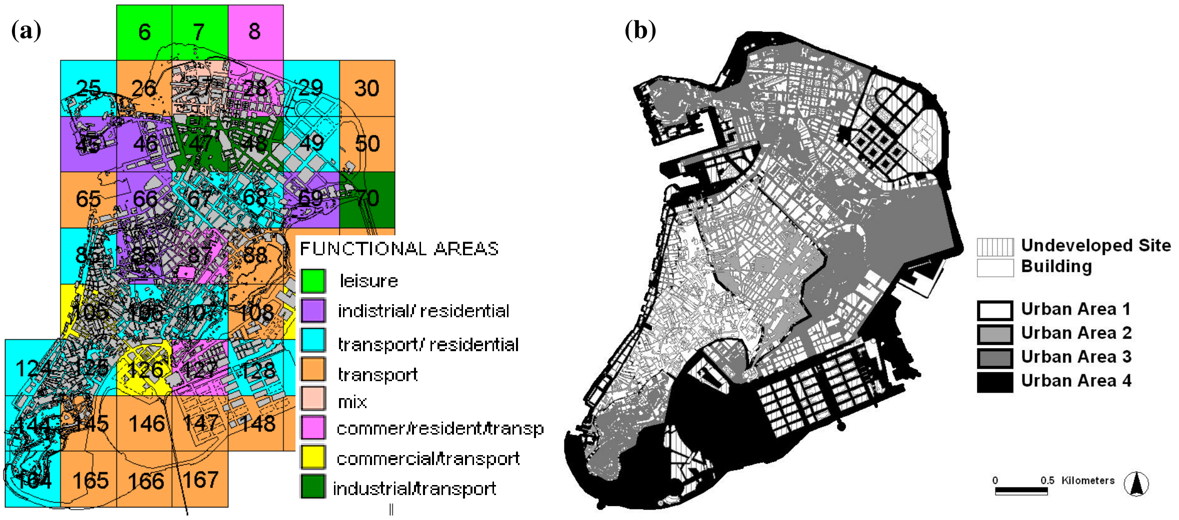


Fig. 5. Urban form studies based on (a) 500×500 m grids (Bai et al., 2002) and (b) urban growth areas.

The neglect of urban geometry in the previous urban forms studies might be due to the involvement of intensive labour costs for manual extraction of street configuration data. In this study, however, the landscape model developed for the building-based approach can extract the complex spatial data of individual buildings from digital maps automatically, so that the intensive labour costs and unintentional human errors due to manual data input and manual operations can be avoided.

With the building-based approach, the influences of urban forms on street environment in Macau are investigated. Four urban forms (i.e., Urban Forms 1-4) located in four associated urban growth areas (i.e., Urban Areas 1-4) in Macau are considered, see Fig. 5b. Urban Areas 1-4 refer to the urban growth areas during four stages, i.e., 1557-1794, 1794-1912, 1912-1957, and 1957-2005 (Tang and Wang, 2007). The street environment in each urban form is assessed in terms of the modeled air and noise pollution in front of the 1343 buildings located in the area of the urban form. Base on the attribute data sets of street configuration data in the receptor point digital map, the road width and building height distributions in each urban form can be summarized quantitatively in Table 1.

In Urban Form 1, the urban morphology is in compact style and the layout of the road network is in complex curvature style for pedestrian transport in the Middle Age. About 65.8% of the roads are less than 10 m wide and 78.8% of the buildings are 10-20 m in height. In Urban Form 2, the proportions of narrow roads less than 10 m wide and low-rise buildings 10-20 m in height decrease, whereas the proportions of roads 10-20 m wide and buildings 20-

30 m in height increase doubly. In Urban Form 3, the proportion of roads less than 20 m wide decreases to 65.5%, while that of roads 20-30 m wide has increased to 23.2%. In Urban Form 4, the proportion of roads 20-30 m wide has increased significantly to 67.7% to serve the increasing demands of vehicle transport. The proportion of high-rise buildings 50-60 m in height has increased to 39.5% to maintain high gross densities.

Road Width (m)	Urban Form 1	Urban Form 2	Urban Form 3	Urban Form 4
0-10	65.8%	48.7%	8.8%	2.9%
10-20	20.5%	40.6%	56.7%	12.3%
20-30	11.7%	7.3%	23.2%	67.7%
30-40	0.6%	1.4%	8.1%	11.7%
40-50	1.4%	2.0%	3.2%	5.3%

Building Height (m)	Urban Form 1	Urban Form 2	Urban Form 3	Urban Form 4
0-10	3.3%	10.7%	15.4%	23.1%
10-20	78.8%	61.4%	35.5%	17.7%
20-30	12.1%	21.6%	16.2%	5.4%
30-40	0.8%	3.3%	4.6%	1.1%
40-50	0.3%	1.7%	6.5%	1.8%
50-60	1.1%	1.3%	9.2%	39.5%
60-70	3.3%	0.0%	7.0%	1.6%
70-80	0.0%	0.0%	3.9%	2.7%
80-90	0.3%	0.0%	1.7%	7.2%

Table 1. Road width and building height distributions in each urban form.

With the building-based approach, the relationship between street configurations, traffic conditions and street environments for each receptor can be investigated. The dependence of road width (canyon width) on aspect ratio (building height : street width) in Macau is shown in Fig. 6a. Since the variations of building heights in Macau are relatively small, the distribution of aspect ratio depends more on the street width. Fig. 6a illustrates a clear pattern that higher aspect ratio appears in narrower street in the peninsula. Such relationship can be described reasonably well ($R^2 = 0.5011$) by the following power model:

$$H/W = 8.1425W^{-0.6938} \tag{1}$$

where H and W are the building height and street width, respectively.

Aspect ratio is a key parameter governing the street-canyon dispersion effect. With higher aspect ratio and decreased dilution of air pollutants from the ground-based traffic emission sources, CO concentrations increase significantly in narrow streets under Traffic Scenario 2 when the spatial variation of traffic condition is neglected, see Fig. 6b. Therefore, CO concentrations are found to be higher in historic areas (Urban Areas 1 and 2) with higher percentage (> 48.7%) of narrow streets (less than 10 m wide) in Traffic Scenario 2, see Fig. 6c. On the other hand, CO concentrations in historic areas are lower when the spatial variation of traffic condition is taken into account in Traffic Scenario 1, i.e., the real traffic condition. The reason is that the traffic capacity, accessibility, and mobility reduce in narrow streets with limited space and complicated morphology in historic areas with Urban Forms 1 and 2, see Fig. 5b. This in turn results in a significant decrease of traffic volumes and traffic emissions in narrow streets which help to maintain the air quality in historic areas. The influences of street configuration on traffic noise distribution and the influences of land use patterns on traffic composition are also observed. Details are discussed in other papers.

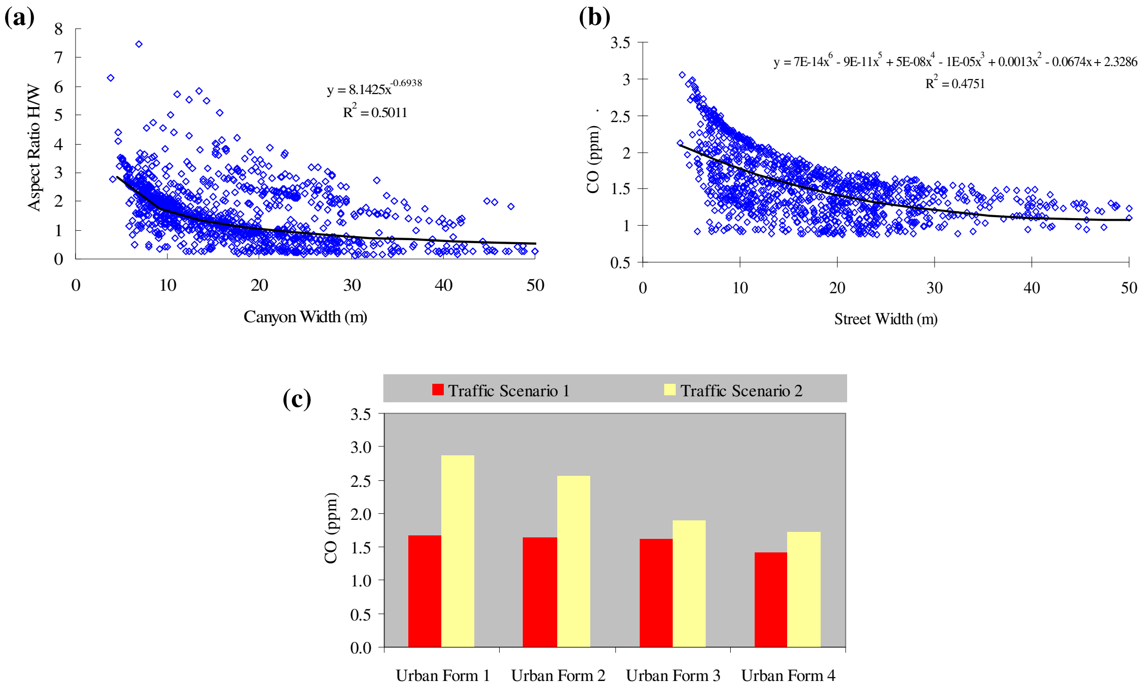


Fig.6. (a) Influence of road width on aspect ratio, (b) Influence of road width on CO concentrations in Traffic Scenario 2, and (c) Influences of urban forms on average CO concentrations.

6. CONCLUSIONS

The data capture and data mining of urban air pollution by an innovated building-based approach are introduced in the paper. Compared with the traditional grid-based approach with spatial resolutions of 1-2 km, the present building-based approach can predict the complex spatial variation of traffic emission, urban geometry, dispersion and air pollution. In the case study in Macau with only 7.7 km^2 of land area, the street configuration data (and air pollutant concentrations) in front of 1343 receptors have been extracted (and modeled) automatically.

Based on the high spatial resolution of input/output data, the complex traffic emissions, the complex street configurations and the non-homogeneous distributions of air quality have been studied. This study shows that the building-based approach may open an innovated methodology in data mining of urban spatial data for environmental assessment.

ABOUT THE AUTHOR: U WA TANG

Contact Details:

Postage address: Faculty of Science and Technology,
University of Macau, Taipa, Macau, PR China.
Phone No: (853) 83974297
E-mail: uwtang@umac.mo; uwtang@gmail.com
Web: http://www.fst.umac.mo/en/staff/fstuwt.html

Education:

B.Sc., M.Sc. & **Ph.D**. in Civil Engineering, University of Macau, Macau

Environmental Research Experience:

- Principal Investigator, Physicochemical Characterization and Biological Activities of Aerosols in Macao, sponsored by the Macao Science and Technology Development Fund (2006-2009)
- Researcher, Characterization and Control of Vehicular Exhaustion Emission Pollution in Macau, sponsored by National Natural Science Foundation of China and University of Macau (2000-2002)
- Researcher, Probabilistic Analysis of Uncertainty in Water Color Remote Sensing Inversion, sponsored by National Natural Science Foundation of China and University of Macau (2000-2002)
- Researcher, Formation and Impacts of Highly Risky Pollution Areas of Non-degradable Organic Species in Pearl River Estuary, sponsored by National Natural Science Foundation of China and University of Macau (2000-2002)
- Researcher, Application of Remote Sensing Technique to Environmental Assessment of Pearl River Estuary, sponsored by University of Macau, Macau Foundation and Chinese National Natural Science Foundation (1996-1999).
- Technical Researcher, 21th Century Macau City Planning Guideline Study (Project Directors: José Chui Sai Ping, Zhao Bing Shi). Fundação para a Cooperação e o Desenvolvimento de Macau, Macao (1999)

Publications:

Journal Papers with SCI International Citation

1. **Tang UW**, Wang ZS, 2007. Influences of urban forms on traffic-induced noise and air pollution: results from a modeling system. *Environmental Modeling and Software* 22: 1750–1764.
2. **Tang UW**, Wang ZS, 2006. Determining gaseous emission factors and driver's particle exposures during traffic congestion by vehicle-following measurement techniques. *Journal of Air and Waste Management Association* 56: 1532–1539.
3. Fu JM, Mai BX, Sheng GY, Zhang G, Wang XM, Peng PA, Xiao XM, Ran R, Cheng FZ, Peng XZ, Wang ZS, **Tang UW**, 2003. Persistent organic pollutants in environment of the Pearl River Delta, China: An overview. *Chemosphere* 52: 1411–1422.
4. Wu Y, Hao JM, Fu LX, Hu JN, Wang ZS, **Tang UW**, 2003. Chemical characteristics of airborne particulate matter near major roads and at background locations in Macao, China. *Science of the Total Environment* 317: 159–172.
5. Wu Y, Hao JM, Fu LX, Wang ZS, **Tang UW**, 2002. Vertical and horizontal profiles of airborne particulate matter near major roads in Macao, China. *Atmospheric Environment* 36: 4907–4918.
6. Qi SH, Yan J, Zhang G, Fu J, Sheng G, Wang ZS, Tong SM, **Tang UW**, Min YS, 2001. Distribution of polycyclic aromatic hydrocarbons in aerosols and dustfall in Macao. *Environmental Monitoring and Assessment* 72: 115–127.
7. He DQ, Zhou ZP, He KB, Liu Y, Hao JM, Wang ZS, **Tang UW**, 2000. Assessment of traffic related air pollution in urban areas of Macao. *Journal of Environmental Sciences* 12(1): 39–46.

Journal Papers with EI International Citation

1. Hao YM, Mai BX, Li CL, Sheng GY, Fu JM, Wang ZS, **Tang UW**, 2004. Primarily study of alkylphenols in sediment core from Nam Van artificial lake of Macao. *Environmental Science* 25(4): 129-33 (in Chinese with abstract in English).
2. Hu JN, Hao JM, Fu LX, Wu Y, Wang ZS, **Tang UW**, 2004. Study on on-board measurements and modeling of vehicular emissions. *Environmental Science* 25: 19–25 (in Chinese with abstract in English).
3. Wang JP, Cheng ST, Jia HF, Wang ZS, **Tang UW**, 2003. An artificial neural network model for lake color inversion using TM imagery. *Environmental Science* 24(2): 73-76 (in Chinese with abstract in English).
4. Wu Y, Hao JM, Fu LX, Hu JN, Wang ZS, **Tang UW**, 2002. Emission inventory for mobile sources in Macao, China. *J Tsinghua Univ* (Sci & Tech) 42(12): 1601–1604 (in Chinese with abstract in English).
5. Qi SH, Fu JM, Sheng GY, Wang ZS, Tong SM, Min YS, Zhang G, **Tang UW**, 2001. Study on priority polycyclic aromatic hydrocarbons (PAHs) in dustfall at Macao. *Research of Environmental Sciences* 14(1): 9–13 (in Chinese with abstract in English).
6. Qi SH, Fu JM, Sheng GY, Wang ZS, Tong SM, Min YS, Zhang G, **Tang UW**, 2000. Study on polycyclic aromatic hydrocarbons in aerosols at Macao. *Research of Environmental Sciences* 13(4): 9–13 (in Chinese with abstract in English).

REFERENCES

Alberti, M., Waddell, P., 2000. An integrated urban development and ecological simulation model. Integrated Assessment 1(3): 215-227.

Bai, S.T., Fu, L.X., Wang, Z.S., Tam, W.M., Chan, H.I., Tang, U.W., 2002. Research of noise simulation and control in Macao, In: Wang ZS, Chen JN, Du PF (Eds.), Investigação Científica Sobre O Ambiente E Desenvolvimento Urbano Em Macau, pp.100–106.

Blair, J., Johnson, K., Carruthers, D., 2003. Modelling Air Quality for London Using ADMS-Urban. Cambridge Environmental Research Consultants TR-0314.

Bøhler, T., Sivertsen, B., 1998. A Modern Air Quality Management System Used in Norway. Kjeller, Norwegian Institute for Air Research (NILU F 4/98).

Borrego, C., Tchepel, O., Barros, N., Miranda, A.I., 2000. Impact of road traffic emissions on air quality of the Lisbon region. Atmospheric Environment 34 (27), 4683-4690.

Borrego, C., Tchepel, O., Costa, A.M., Amorim, J.H., Miranda A.I., 2003. Emission and dispersion modelling of Lisbon air quality at local scale. Atmospheric Environment 37 (37), 5197-5205.

Borrego, C., Tchepel, O., Salmim, L., Amorim, J.H., Costa, A.M., 2004. Integrated modelling of road traffic emissions: application to Lisbon air quality management. Cybernetics and Systems 35, 535–548.

Borrego, C., Martins H, Tchepel O, Salmim L, Monteiro A, Miranda AI, 2006. How urban structure can affect city sustainability from an air quality perspective. Environmental Modelling and Software 21 (4): 461–467.

Chen, C., Wang XM, Zhang G, Min YS, Sheng GY, Fu JM, Wang ZS, Tang UW, 2000. Volatile Organic Compounds in Ambient Air in Macau. In: Wang ZS (Ed.), Environmental Assessment of Vehicular Exhaustion Emission Pollution in Urban Areas of Macau, University of Macau, Macao, pp.32-39 (in Chinese with Abstract in English).

De Ridder, K., Adamec, V., Bañuelos, A., Bruse, M., Bürger, M., Damsgaard, O., Dufek, J., Hirsch, J., Lefebre, F., Pérez-Lacorzana, J.M., Thierry, A., Weber, C., 2004. Short communication: an integrated methodology to assess the benefits of urban green space. Science of the Total Environment 334–335 (1), 489–497.

Department of the Environment, 1992. Development Plans and Regional Planning Guidance. PPG12. HMSO, London.

Department of Transport, Welsh Office, 1988. Calculation of Road Traffic Noise. HMSO: London.

Gontier, M., 2007. Scale issues in the assessment of ecological impacts using a GIS-based habitat model — A case study for the Stockholm region. Environmental Impact Assessment Review 27, 440–459.

Goodchild, M.F., Quattrochi, D.A., 1997. Introduction: scale, multiscaling, remote sensing, and GIS. In: Quattrochi DA, Goodchild MF, editors. Scale in remote sensing and GIS. Boca Raton, FL: Lewis Publishers; 1997. 1 –11.

Hao, J.M., Hu, J.N., Fu, L.X., Wu, Y., Wang, L.T., 2003. Vehicle emission characteristics and control in Macao. In: Wang, Z.S. (Ed.), Scientific Innovation for Environmental Assessment of Coastal Cities. University of Macau, Macao, pp. 67–130.

Hertel, O. and Berkowicz, R., 1989. Operational Street Pollution Model (OSPM), evaluation of model on data from St. Olavs Street in Oslo. DMU Luft A-135. National Environmental Research Institute, Roskilde, 34p.

ITT – Institut fur Technische Thermodynamik, 1994. The nonhydrostatic mesoscale model MEMO: version 5.0 – A technical reference. Karlsruhe.

Jensen, S.S., 1998. Mapping human exposure to traffic air pollution using GIS. Journal of Hazardous Materials 61 (1–3), 385–392.

Jensen, S.S., 1999. A geographic approach to modelling human exposure to traffic air pollution using GIS. National Environmental Research Institute, Roskilde.

João, E., 2002. How scale affects environmental impact assessment. Environ Impact Asses Rev, Vol. 22(4), 287–306.

Kadikis, N., Vasiljeva, T., Jevtushenko, L., Lyulko, I., Smalins, E., 2002. Air quality modelling as a part of complex air management system in Latvia. Proceedings of the 8th International Conference on Harmonisation within Atmospheric Dispersion Modelling for Regulatory Purposes.

Marquez, L.O., Smith, N.C., 1999. A framework for linking urban form and air quality. Environmental Modelling and Software 14, 541–548.

McHugh, C.A., Carruthers, D.J., Edmunds, H.A., 1997. ADMS and ADMS-Urban. International Journal of Environment and Pollution 8(3–6), 438–440.

Montello, D.R., Golledge, R., 1999. Scale and detail in the Cognition of Geographic Information—report of a specialist meeting held under the auspices of the Varenius Project. Santa Barbara, CA: National Centre for Geographical Information and Analysis.

Moussiopoulos, N., 1995. MARS: Version 2.0 – User's Guide. Aristotle University, Thessaloniki.

National Physical Planning Agency, 1991. Summary of the Fourth Report Extra on Physical Planning. Ministry of Housing, Physical Planning and Environment, the Hague, the Netherlands.

Newton, P.W., 1997. Re-shaping Cities for a More Sustainable Future. Australian Academy of Technological Sciences and Engineering, Canberra (also AHURI, Melbourne).

Partidário, M.R., 2007. Scales and associated data — What is enough for SEA needs? Environmental Impact Assessment Review 27 (2007) 460–478.

Qi, S.H., Yan, J., Zhang, G., Fu, J., Sheng, G., Wang, Z.S., Tong, S.M., Tang, U.W., Min, Y.S., 2001a. Distribution of polycyclic aromatic hydrocarbons in aerosols and dustfall in Macao. Environmental Monitoring and Assessment 72: 115–127.

Qi, S.H., Fu, J.M., Sheng, G.Y., Wang, Z.S., Tong, S.M., Min, Y.S., Zhang, G., Tang, U.W., 2001b. Study on priority polycyclic aromatic hydrocarbons (PAHs) in dustfall at Macao. Research of Environmental Sciences 14(1): 9–13 (in Chinese with abstract in English).

Tang, U.W., Wang, Z.S., 2006. Determining gaseous emission factors and driver's particle exposures during traffic congestion by vehicle-following measurement techniques. Journal of Air and Waste Management Association 56, 1532–1539.

Tang, U.W., Wang, Z.S., 2007. Influences of urban forms on traffic-induced noise and air pollution: results from a modelling system. Environmental Modelling and Software 22: 1750–1764.

Wu, Y., Hao, J.M., Fu, L.X., Wang, Z.S., Tang, U.W., 2002a. Vertical and horizontal profiles of airborne particulate matter near major roads in Macao, China. Atmospheric Environment 36: 4907–4918.

Wu, Y., Hao, J.M., Fu, L.X., Hu, J.N., Wang, Z.S., Tang, U.W., 2002b. Emission inventory for mobile sources in Macao, China. J Tsinghua Univ (Sci & Tech) 42(12): 1601–1604 (in Chinese with abstract in English).