Title
A GIS-Based Approach to Spatial Allocation of Area Source Solvent Emissions

Journal
Environmental Modelling and Software, 15

Authors
Dai, Jian
Rocke, David

Publication Date
2000

Peer reviewed

ELSEVIER

Environmental Modelling & Software 15 (2000) 293–302

Environmental
Modelling & Software

www.elsevier.com/locate/envsoft

A GIS-based approach to spatial allocation of area source solvent emissions

Jian Dai, David M. Rocke

Center for Image Processing and Integrated Computing (CIPIC), University of California, Davis, One Shields Avenue, Davis, CA 95616-8553, USA

Accepted 15 July 1999

Abstract

Current area source emission inventories estimate total emissions for various industrial, commercial, and consumer activities at the county or higher levels. The lack of emission estimates at subcounty levels severely limits the modeling and planning capabilities in urban and regional air quality management. This paper extends current area source emission inventory methodology by developing a geographical information system (GIS) based approach for allocating county-level emission estimates to subcounty units. The new methodology uses GIS to develop and integrate spatial data, to analyze spatial variations in emissions, and to derive input to cell-based air pollution models. This approach incorporates statistical models to predict the spatial distribution of emission source activities based on widely available data. The paper presents an application of the proposed approach to emission inventory of the adhesives and sealants category in the Sacramento modeling region, California.

Keywords: Area source emissions; Spatial modeling; Geographic information systems

Software availability
Name of Software: AML tools for data processing
Developer: J. Dai and D.M. Rocke, CIPIC, University of California, Davis, CA 95616
Year first available: 1997
Software required: Arc/Info GIS
Program language: Arc/Info AML
Availability and cost: available at cost of transfer

1. Introduction

In emission inventory, area sources are generally defined as those sources that individually emit relatively small quantities of air pollutants but collectively result in significant emissions (CARB, 1995). Area source inventories estimate emissions by areas for various industrial, commercial and consumer activities. In the United States, the current inventory methods for area sources are generally derived from a basic methodology developed in the 1970s as a part of the National Emissions Data System (NEDS), which uses a material balance to quantify national solvent usage and allocates national estimates to States and counties (USEPA, 1975; Battye et al., 1993; CARB, 1995). Those methods are easy to use but lack the ability to integrate and analyze spatial information and thus cannot be used effectively to derive emission estimates at subcounty levels. Information on spatial distributions of emissions within a county is however important in air quality management for at least two reasons. First, it helps identify local areas of pollution concentration where special measures may be needed to reduce pollution. Second, it provides necessary input to air quality simulation models such as the Urban Airshed Model (Morris and Myers, 1990; Rao, 1987; Scheffe, 1990), which require data on emission distributions in subcounty modeling units. The Urban Airshed Model (UAM) is a three-dimensional photochemical grid model designed to calculate the concentrations of air pollutants by simulating the physical and chemical processes in the atmosphere that affect pol-

* Corresponding author. Tel.: 530-752-0510; fax: 530-752-8894.
E-mail addresses: jjdai@ucdavis.edu (J. Dai), dmrocke@ucdavis.edu (D.M. Rocke).

1364-8152/00/$ - see front matter
PII: S1364-8152(00)00004-9

lutant concentration. The purpose of this paper is to extend the current methodology for area source emission inventory by developing a geographical information system (GIS) based approach for analyzing small area variations in emission source activities and allocating county-level emission estimates to subcounty units (Fig. 1).

A GIS has been an important tool for environmental modeling (e.g. Goodchild et al. 1993, 1996; USEPA, 1995) and air pollution analysis (Souleyrette, 1992; Shimp and Campbell, 1995; Bachman et al., 1996; Jensen and Sathisan, 1996). GIS is often used as the integrating framework for spatial data analysis and modeling (e.g. Longley and Batty, 1996; USEPA, 1998; Brodie, 1999). This study uses GIS to develop and integrate spatial data, to analyze spatial variations in source activities, and to derive input to cell-based air quality models. The procedure has three main steps. The first step is to develop a spatial database for emission allocation. In this step, activities that generate specific emissions are identified and their locations are geocoded; data measuring the activities and characterizing the area are collected and processed. The second step is to disaggregate the study area into subcounty units (e.g. modeling cells) and convert all data into the modeling unit. This is done by spatial overlay operations. In the third step, spatial variations in the activities are mapped, evaluated, and modeled; and emission allocation factors are computed. These factors will be used to allocate county-level estimates to the modeling units. Statistical models can be used in this step to predict the spatial pattern of activities, thereby facilitating air quality planning. The methods will be illustrated with an application to emission inventory of the adhesives and sealants category in the Sacramento modeling region, California.

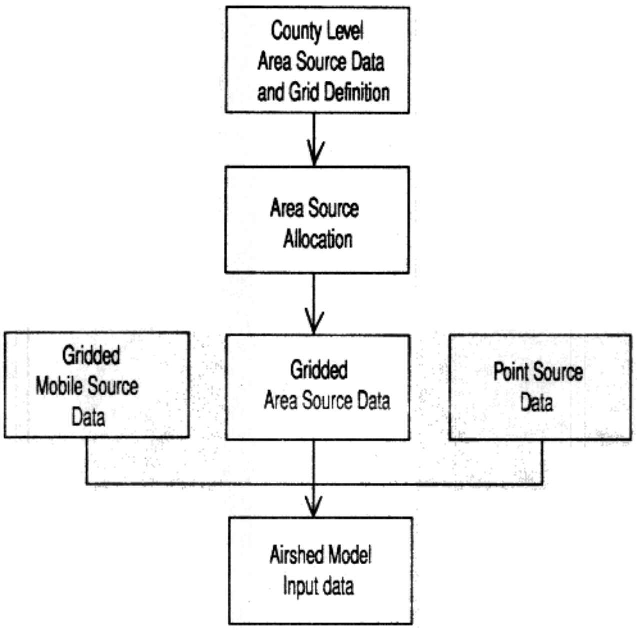

Fig. Emission data input to airshed models.

The paper is organized as follows. Section 2 discusses the spatial analysis methodology, including spatial database development, map overlay, spatial allocation, and spatial pattern modeling. Section 3 presents an application of the methods to spatial allocation of emission estimates for a particular source category, namely adhesives and sealants. The empirical data is collected from the Sacramento modeling region in California. Section 4 provides concluding remarks.

2. Methodology

Emission inventories can be developed by identifying the sources of pollution, measuring the extent of the polluting process, and determining polluting potential. Area sources are aggregates of emission sources that are too small or too numerous to be included individually in a point source inventory. Thus area source emissions are usually measured by activity statistics. In the standard NEDS method for area source solvent emissions, inventories are made by allocation of national solvent usage to the States and counties based on census data on employment, population, or other surrogate parameters. To obtain sub-county small-area emission estimates such as those required by cell-based air pollution models the standard method has to be extended. In the new method, we use spatial intensity of the sources as the measurement to allocate county-level emissions to subcounty units. For a given source category, we identify the location of sources activities, measure and model the spatial intensities (e.g. number of emission sources per unit of area), and obtain allocation factors for the modeling areas. This approach entails intensive spatial data operations. Hence, a set of spatial data processing and analysis tools are required. For example, geocoding tools will be used for mapping the location of source activities; spatial overlay tools will be needed for integrating spatial data that may be in different units; and spatial modeling tools will be employed to estimate source activity patterns. This section describes the methods used in this study.

2.1. Spatial database development

The basic spatial objects in a GIS database are points, lines, and polygons (areas). In the database, the location of a pollution source such as an industrial facility is represented by a point entity consisting of XY coordinates and some attributes. The location of a facility can be identified by its address available from a variety of sources such as phone books, commercial business lists, government records, or online databases. The source location layer is created using geocoding methods. Geocoding is a mechanism for building a database relationship between addresses and spatial features. In geocoding, a GIS compares the address of a facility against those on a street layer (such as the U.S. Census Bureau's TIGER files) in the database. When a match is found, a geographic coordinate is calculated for the address, and a spatial point representing the source location is created in the database. If the address isn't matched, the GIS would display diagnostic messages that explain why a match was not found, and allow the address in question to be edited and the matching process to be restarted. Specifically, the Arc/Info GIS (ESRI, 1997) provides the following geocoding capabilities: creating the address coverage (or converting TIGER files to an address coverage); building and maintaining INFO (a relational database) files containing a list of addresses to be matched; matching the list to the address coverage to create points (or matching addresses interactively to specified locations); processing unmatched addresses; and maintaining the address coverage.

Selection of the areal unit for allocation depends on the purpose of the study. For air quality modeling, grid cells are usually chosen because photochemical air quality simulation models such as the Urban Airshed Model use grid cells as the areal unit. For the Urban Airshed Model, the cell size is usually in the range of 2 to 10 km (Morris and Myers, 1990). The grid layer can be created easily using GIS's macro language capabilities. First, the coordinates for the intersections of the grid lines are computed. Then, using the coordinates, grid squares are created and built into polygon topology.

The database usually also contains other data sets relevant to the source activities or characterizing the study. For example, it is often useful to show where the emission sources are located in relation to land uses, road network, and population distribution. These data are also useful in modeling the spatial pattern of source activities.

2.2. Spatial data overlay and spatial allocation

2.2.1. Spatial overlay

The spatial database for emission analysis may consist of a variety of elements in different spatial units and at a variety of resolutions. For example, we code the emission sources as spatial points but emissions should be allocated to areal units. Demographic data are usually available in census units such as census tracts or census blocks, whereas urban transportation data are often in traffic zones. It is a common problem in environmental modeling that the spatial units for which data are available are not necessarily the same as the one that the analysis requires. For example, the unit for air quality modeling may be a grid cell, but available data may be in a variety of other units. A solution to the problem is spatial overlay. A GIS has strong spatial overlay capabilities and therefore is ideally suited to derive data in the target unit given the relevant source data. The following spatial overlays are often needed in preparing data for cell-based air quality modeling:

- point in polygon operation (e.g. count the number of emission source points in each cell);
- polygon on polygon operation (e.g. convert from census units to cell-based units); and
- line in polygon operation (e.g. compute the length of a linear feature such as highway within a cell).

Although the application of spatial overlay methods is straightforward, caution must be taken when polygon overlays are involved. A technical issue in polygon overlay is known as the areal weighting problem (Goodchild and Lam, 1980). Since boundaries of a source zone (polygon) usually do not coincide with those of a target zone (polygon), one must weight the source zone values according to the area of the target zone they make up. The method is called the areal weighting method and is discussed below. An example of polygon overlay is given in Fig. 2.

Denote V the variable of interest, S the source zone, T the target zone, and A the area of a zone. Suppose, for a given area, we want to convert the population density in census tracts to those in modeling cells. In this case, V is the population density variable, S is a census tract, and T is a grid cell. As S intersects T, their boundaries form a zone of intersection ST. The problem is to find the value of V for the target zone or the intersection zone. The computation depends on the measurement of

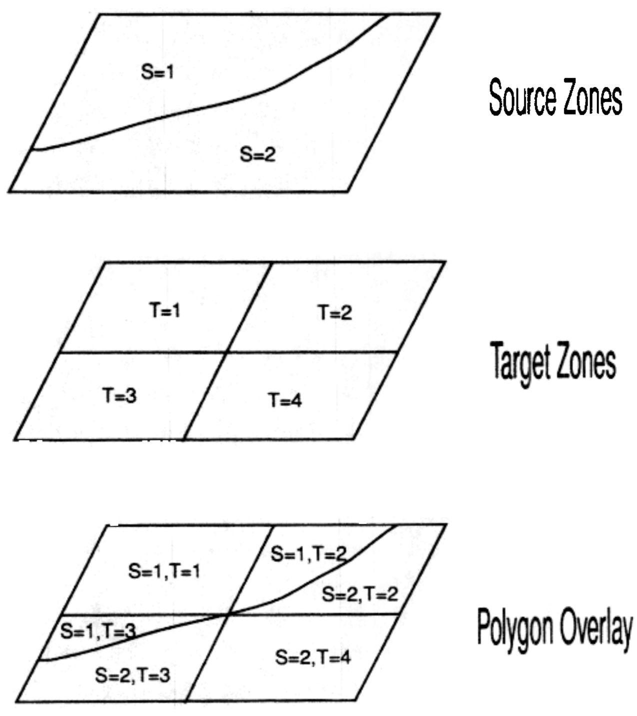

Fig. 2. Example of polygon overlay.

V, that is, whether it is "extensive" or "intensive" (Goodchild and Lam, 1980). Variable V is extensive if its value for a target zone is equal to the sum of its values for the intersection zone. V is intensive if its value for a target zone is the weighted average of its value for the intersection zones. V is usually considered to be extensive if it is a count variable (e.g. number of people in a census tract), or intensive when it is a proportion, percentage or rate (e.g. population density).

Assuming that V is evenly distributed within the source zone, the values of V can be computed as follows (Flowerdew and Green, 1994). If V is an extensive variable, its values are given by

$$V_t=\sum_s \frac{V_s A_{st}}{A_s}, \qquad (1)$$

Or, if V is an intensive variable, the value is computed as

$$V_t=\sum_s \frac{V_s A_{st}}{A_t} \qquad (2)$$

The assumption that the variable of interest is uniformly distributed over the source zone is not always plausible. For example, there might be a lake in the zone. In those cases, the methods of areal interpolation using ancillary data (Green, 1990) may be used which take into account other relevant information available about the source zones.

2.2.2. Spatial allocation

To allocate county-level emission estimates to sub-county units, we use a measure of spatial intensity selected for the source category. For example, if the source category is automobile refinishing, the measure could be the number of automobile refinishing shops (number of the emission sources) or number of employees in the automobile refinishing business (size of the activities) per unit area. Selection of the measure depends on the purpose of the study and the availability of data. The spatial intensity is computed by GIS using the information on location and attributes of the sources. The allocation factor for an areal unit is obtained by dividing the measure for the unit by the county total.

Two methods of spatial allocation are considered here. The first method may be called discrete allocation, which treats the study area as a collection of discrete units. Assuming that grid cells are the objects of allocation, the source point layer needs to be overlaid with the grid layer and then allocation factors for grid cells are computed. Let a_{ij} be the data value of point j (j=1,…,M) in cell i (i=1,…,N). The allocation factor for cell i, w_i, is computed by

$$w_i=\frac{\sum_{j=1}^{M} a_{ij}}{\sum_{i=1}^{N}\sum_{j=1}^{M} a_{ij}} \qquad (3)$$

The second method, termed surface allocation, considers the study area as a smooth surface. In this method, an activity surface is first interpolated using data values of the source points. Then the surface is overlaid with the grid layer, and the spatial intensity of each cell is obtained by averaging the surface values within the cell. Finally allocation factors are computed. The surface can be generated by a GIS using several methods, the most common one being the Triangulated Irregular Network (TIN) data model (Peuker et al., 1978; ESRI, 1996). The TIN is a surface model that uses a sheet of continuous, connected triangular facets based on a Delaunay triangulation of irregularly spaced sample points. Mass points, breaklines and exclusion polygons control the behavior of TIN surface-modeling operator. In applications, many details must be considered. For example, one should determine if the sample points adequately capture the maxima and minima of the surface and if it is reasonable to use linear interpolation. Surface interpolation is often an iterative process (Brodie, 1999) where data values generated from the TIN need to be verified against the original data for goodness of fit. The sample points can then be changed and a new TIN is generated. The process goes on until the surface is adequately represented. Using either the discrete or surface allocation method, allocation factors can be calculated. In applications, spatial overlay operations and spatial allocations can be automated using the GIS's programming capabilities such as AML available in Arc/Info.

2.3. Modeling spatial patterns

Spatial allocation can be done using either observed or projected data on emissions. Data projection or modeling is necessary when the air quality simulation model (such as the UAM) is used to develop future ozone air quality plans. Modeling also provides a low-cost approach for updating emission inventories. For example, the California Air Resources Board has requested that the emission allocation factors be estimated and updated using easily available data. Our objective of modeling is to estimate a quantitative relationship between the measure of spatial intensity and data that are easily available such as population, land use, and highways. A statistical model is a useful tool for the purpose. The selection of a particular statistical model depends on the measurement of the endogenous variable and the availability of data. A useful model can be made by assuming that the activities follow the Poisson distribution (Dai and Rocke, 1998). The Poisson model is a spatial point process model and has been widely used for spatial analysis.

Let y_i be the number of source activities in a given unit, the Poisson model assumes that each y_i is drawn from a Poisson distribution with parameter λ_i (i =1,..,N)

$$\Pr(Y_i=y_i)=\frac{e^{-\lambda_i}\lambda_i^{y_i}}{y_i!},\ y_i=0,1,2,... \tag{4}$$

The value of λ indicates the spatial intensity of the phenomenon under study, for example, the expected number of emission-producing facilities per unit of area. To model the variation in spatial intensity, λ is related to a set of predictors, x, by

$$E(y_i|x_i)=\lambda_i=\exp\left(\sum_j \beta_j x_{ij}\right), \tag{5}$$

where β is a set of parameters to be estimated. Eq. (5) gives the Poisson regression model, which can be used to predict the distribution of source activities across the study area. The Poisson regression model is a member of the generalized linear model (GLM) family (McCullagh and Nelder, 1989) and has been widely used to model spatially distributed events. The model's assumption of equal mean and variance of y is however rather restrictive. In applications, this assumption may not be always held. For example, if the count data contains many zeros, the Poisson variance tends to be larger than its mean. This can happen in developing emission estimates as the modeling cells covering rural areas seldom contain emission sources. To deal with this extra-Poisson variation, we use the variance function

$$Var(y_i|x_i)=\sigma^2\lambda(x_i), \tag{6}$$

where σ^2 (>1) is called the over-dispersion parameter and can be estimated by (McCullagh and Nelder, 1989; Winkelmann, 1994),

$$\hat{\sigma}^2=\frac{1}{N-P}\sum_{i=1}^{N}\frac{(y_i-\hat{\lambda}_i)^2}{\hat{\lambda}_i} \tag{7}$$

In Eq. (7), N is the sample size and P is the number of regression coefficients in the model. The model's goodness-of-fit can be assessed using the deviance function (McCullagh and Nelder, 1989)

$$D(y;\lambda)=2\sum_{i=1}^{N}\{y_i\ln(y_i|\lambda_i)-(y_i-\lambda_i)\}. \tag{8}$$

In the deviance test the goodness-of-fit is measured by the difference between the maximum log likelihood achievable and the log likelihood achieved by the model under the test. The test statistic has asymptotically a Chi-square distribution with $(N-P)$ degrees of freedom.

Software for estimating Poisson regression models and generalized linear models is readily available, including S-Plus (Chambers and Hastie, 1993; Venables and Ripley, 1987), Stata (Stata, 1998), Gauss (Aptech Systems, 1996), SAS (SAS Institute, 1993), and GLIM (Aitkin et al., 1989).

3. Application

We have applied the proposed methodology to analyze small area variations in emission sources and allocate county-level emission estimates to subcounty units. The emission source category selected for this application is the adhesive and sealant category, which is one of the area source categories inventoried by the California Air Resources Board (CARB, 1995). The main pollutants of this category are the total organic gas (TOG) emissions from solvents contained in adhesives and sealants used by various industries. The study area is the Sacramento modeling region, consisting of three counties, Sacramento, Solano and Yolo, in the State of California. The study area is shown in Fig. 3. The UAM used in the Sacramento State Implementation Plan (SIP) requires the input emission data in small spatial units (modeling cells). However, the emission estimates are available only at the county or even larger areal units. Thus, it is necessary to allocate the county-level estimates to modeling cells. Moreover, as requested by the CARB, we developed statistical models so that spatial variations in source activities could be estimated based on easily available data. The modeling facilitates the development of future air quality plans and reduces the cost of emission inventory updates. The application consists of identifying the emission sources, measuring the spatial intensity of source activities, modeling the spatial variations, and computing allocation factors.

3.1. Database development

The GIS database consists of data layers that show the location of the source activities, that can be used to model the spatial intensity of the source activities (such as employment, population distribution, and land uses) and that contain grid cells for computing cell based data values and allocation factors. First, we had to identify the sources activities, that is, industrial uses of adhesives and sealants in the study area. Although applications for adhesives and sealants span almost the entire range of industries, the key industries using adhesives and sealants are limited to a few. A recent study on inventory database for area source solvent emissions (Battye et al., 1993) indicates that two industries alone, the paper packaging and wood products industries, account for over 80% of total industrial adhesive and sealant solvent use in the US. Other significant users are found in the furni-

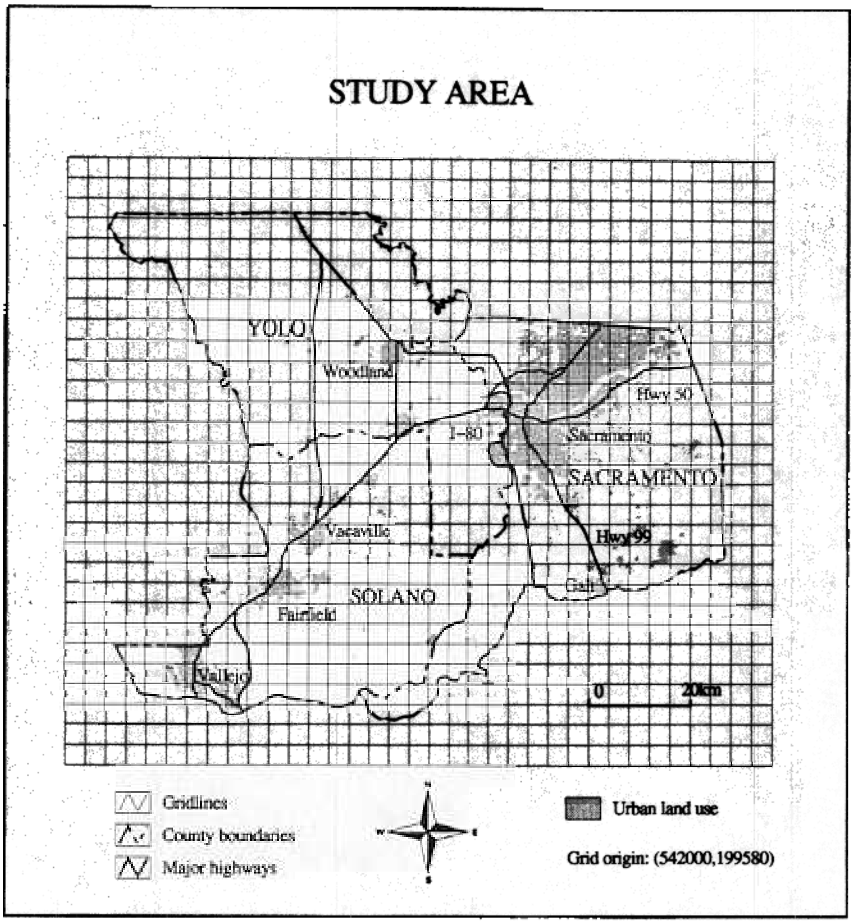

Fig. 3. The study area.

ture and home-building industry and the printing and publishing industry. The businesses in those industries were identified using the standard industrial classification (SIC) codes:

- the wood products industry (SIC=24);
- the furniture and home-building industry (SIC=25);
- the packaging industry (SIC=26); and
- the printing and publishing industry (SIC=27).

We obtained a list of businesses in those industries with addresses from a commercial business database. We deleted small businesses (less than 15 employees) in the printing and publishing industry which are generally not significant users of adhesives and sealants. Thus, our final list contained 312 businesses of the selected industries in the study area. For analysis and modeling, we collected land use data from the California Department of Water Resources, census data on employment and population distribution with census tract boundaries from the Teale Data Center in Sacramento, and the TIGER/Line files from the US Census Bureau. The data came in different formats and projections. We converted all GIS data to the Arc/Info format and the UTM (Universal Transverse Mercator) coordinate system. Source locations were geocoded using the address-matching tool in Arc/Info GIS.

The modeling grid, covering the whole study area, was created as a polygon coverage. It consists of 1050 square cells in an array of 30 rows by 35 columns; with each representing a zone of 4 km by 4 km (see Fig. 3). The 4 km cell size was considered to be appropriate for simulating air quality in the study region. All data layers were integrated with the grid layer to generate cell-based data. Specifically, point in polygon operations were performed to count the number of major industrial users of adhesives and sealants in each cell; line in polygon operations were done to compute the length of highways within each cell; and polygon in polygon operations were carried out to obtain population and employment densities and percentage of urban land use for each cell. The data integration operations were automated using programs written in the Arc Macro Language (AML).

3.2. Mapping and modeling

The study area is located in the Central Valley of California. The three-county area has a total population of

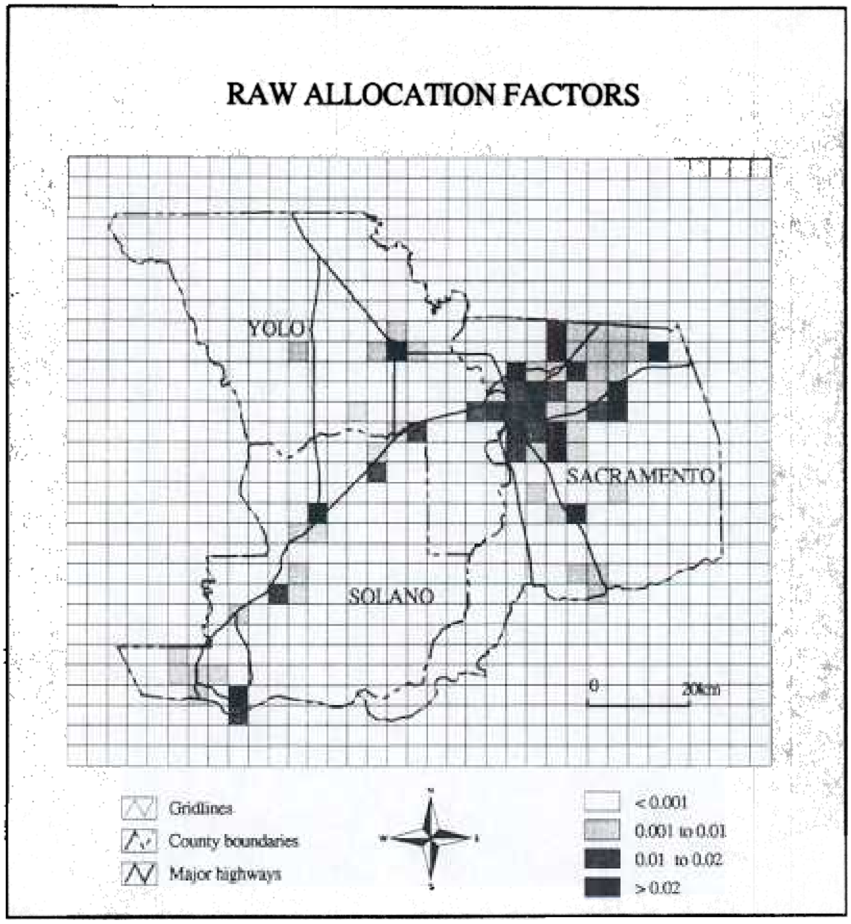

Fig. 4. Raw allocation factors.

1.67 million, about 68% of which is in the Sacramento county. The Sacramento metropolitan area is the dominant urban center in the region. Two interstate highways, I-80 and I-5, go through the region. The industrial users of adhesives and sealants in the study area are mainly located in urban areas and along major highways. Most of the users are found within the Sacramento metropolitan area. In Solano County, the users are mainly concentrated in cities along the interstate highway I-80, such as Vallejo, Fairfield, and Vacaville. Yolo County is the most sparsely populated area in the region. It has fewer urban areas and much fewer businesses in the industries than the other counties. The data show that the distribution of the contributing sources is uneven through the region and is far from homogeneous within each county. The data suggests that some of the areal attributes such as population density, urban land use, and proximity to major highways might be good predictors of the spatial distribution of industrial uses of adhesives and sealants.

Table 1
Definition of variables

Variable name	Definition
Dependent variable	
NSP	Number of source points per cell
Predictors	
URBAN	Percentage of urban land use
URBAN2	URBAN square
HWY	Miles of highway
POP	Population density (1000 persons/square miles)
MFGEMP	Manufacturing employment density (1000 employees/square miles)
POPMFG	POP * MFGEMP

Spatial allocations are done using the measure of spatial intensity of the sources. In this study, the spatial intensity is measured by the number of industrial users per modeling cell. Allocation factors were calculated using Eq. (3). Those are raw allocation factors since they were computed using the raw data. The map of raw allocation factors is shown in Fig. 4. On the map allocation factors are classified into four categories by their values, shown in different colors: the larger the allocation values, the darker the color. Thus, cells with darker colors indicate areas with higher level of source activity, possible areas of concentration of the emissions.

The Poisson regression model was used to model the spatial distribution of the source activities. In the model,

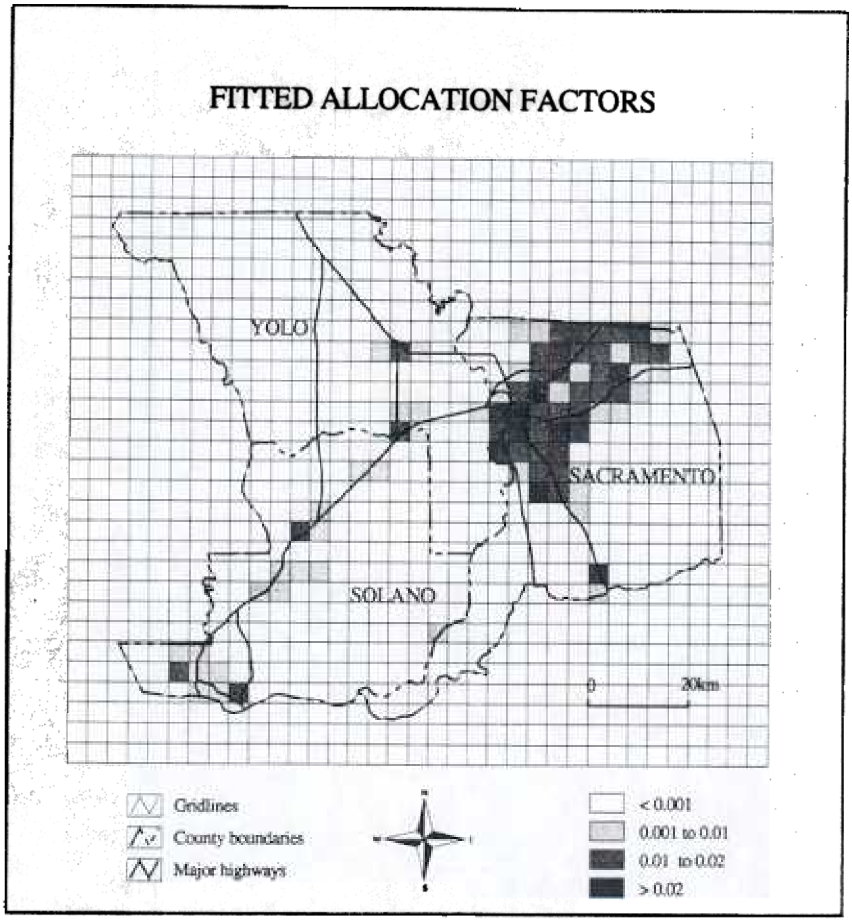

Fig. 5. Fitted allocation factors.

the dependent variable is the number of source points per cell (mean=0.5725, standard deviation=2.3931). The independent variables are percentage of urban land use, miles of highways, population density, and manufacturing employment density. The variables used in the Poisson regression model are defined in Table 1. Since data on those variables are easily available and can be updated on a regular basis, the model is easy to use in practice.

A number of Poisson regression models were estimated. The model fitting the data best is presented in Table 2. The model has five predictors, and all estimated coefficients are statistically significantly different from zero at conventional significance levels. The model's goodness-of-fit can be evaluated using the deviance function defined in Eq. (8). The estimated value is 431.552 with 539 degrees of freedom. This value does not exceed the upper 95% critical point of the Chi-square distribution, indicating no evidence of lack of fit of the data. The estimated value of the overdispersion parameter is 1.9189, greater than one. Hence, the variance of the Poisson variable is "overdispersed". As a result, the standard errors of the regression coefficients must be adjusted upward proportional to the degrees of overdispersion in the data. This was done by dividing the values of the standard errors by the square root of 1.9189. The standard errors shown in Table 2 have been corrected for overdispersion.

Results of the Poisson regression model were used to estimate the expected spatial intensity of the emission sources per cell. Then the fitted allocation factors were

Table 2
Estimated Poisson regression model

Poisson regression		Number of obs	=545
Goodness-of-fit chi2(539)	=431.552	Model chi2 (5)	=1039.273
Prob > Chi2	=0.582	Prob > chi2	=0.000
Log Likelihood	= −369.988	Pseud0 R2	=0.584

Variable	Coefficient	Std. err	Asy. T ratio	P|>t|
URBAN	0.1625	0.0149	10.914	0.000
URBAN2	−0.0012	0.0001	−8.254	0.000
HWY	0.0891	0.0286	3.110	0.001
MFGEMP	−13.0181	4.0127	−3.244	0.001
POPMFG	1.6054	0.6498	2.471	0.007
CONS	−3.1820	0.2741	−11.611	0.000

computed using the discrete allocation method. As before, the allocation values were categorized and a map was made. The map of fitted allocation factors is presented in Fig. 5. A comparison of the fitted map with the raw map shows that the patterns of spatial distribution are similar on the two maps, with the fitted map having a little smooth effect. The estimated model fits the data reasonably well.

4. Conclusions

Current area source emission inventories estimate total emissions for various industrial, commercial, and consumer activities at the county or higher levels. The lack of emission estimates at subcounty levels severely limits the modeling and planning capabilities of urban and regional air quality management. This paper extends current area source emission inventory methodology by developing a GIS-based approach for analyzing small area variations in source activities and allocating county-level emission estimates to subcounty units. This is done by first estimating the spatial intensity of source activities and then using the intensity measure to derive allocation factors. The approach also incorporates statistical models to predict the spatial distribution of source activities using widely available data. The use of statistical models has increased the analytical power of the approach and facilitates emission inventory updates.

The usefulness and flexibility of the proposed approach have been illustrated in the application. The adhesive and sealant category is one of the source categories for which there was a marked lack of available data for emission inventory at subcounty levels. The application consists of identification of the source activities, analysis of the spatial distributions, modeling of the spatial intensities, and estimation of cell-based allocation factors. The data used in the study are easily available and can be updated on a regular basis.

A GIS is an indispensable tool for emission inventory and air quality management. In addition to its powerful capabilities in processing, analyzing and modeling spatial data, GIS provides excellent visualization tools that can be used effectively to present emission inventory results and alternative solutions to air pollution problems. Therefore, the use of GIS can improve not only the analytical capabilities for air quality management but also our ability to communicate work results and research findings to the decision makers and the public in general.

Acknowledgements

The research reported in this paper was supported by grants from the California Air Resources Board and the National Science Foundation (DMS 95-10511, DMS 98-70172, and ACI 96-19020).

References

Aitkin, M., Anderson, D., Francis, B., Hinde, J., 1989. Statistical Modeling in GLIM. Oxford University Press, Oxford, England.

Aptech Systems, 1996. GAUSS: Mathematical and Statistical System. Aptech Systems, Inc, Maple Valley, WA.

Bachman, W., Sarasua, W., Guensler, R., 1996. A GIS Framework for Mobile Emissions Modeling. Transportation Research Board, Washington DC.

Battye, W., Viconovic, G., Williams, A., 1993. Inventory Data Base Analysis for Area Source Solvent Emissions. Report prepared for the U.S. Environmental Protection Agency (EPA 68-D2-0181) by EC/R Incorporated, Durham, North Carolina.

Brodie, R.S., 1999. Integrating GIS and RDBMS technology during construction of a regional groundwater model. Environmental Modeling and Software 14, 119–128.

CARB, 1995. Emission Inventory Procedural Manual, vol. III, Methods for Assessing Area Source Emissions. California Air Resources Board, Sacramento, California.

Chambers, J., Hastie, T., 1993. Statistical Models in S. Chapman and Hall, London, England.

Dai, J., Rocke, D.M., 1998. Modeling Spatial Variation in Area Source Emissions. Working Paper, Center for Image Processing and Integrated Computing, University of California, Davis, California.

ESRI, 1996. Surface Modeling with TIN. Environmental Systems Research Institute, Redlands, California.

ESRI, 1997. ARC/INFO User's Mannual. Environmental Systems Research Institute, Redlands, California.

Flowerdew, R., Green, M., 1994. Areal interpolation and types of data. In: Fotheringham, S., Rogerson, M. (Eds.), Spatial Analysis and GIS. Taylor and Francis Inc, Bristol, PA.

Goodchild, M.F., Lam, N., 1980. Areal interpolation: a variant of the traditional spatial problem. Geo-Processing 1, 297–312.

Goodchild, M.F., Parks, B.O., Steyaert, L.T., 1993. Environmental Modeling with GIS. Oxford University Press, New York.

Goodchild, M.F., Steyaert, L.T., Parks, B.O., Johnston, C., Maidment, D., Crane, M., Glendinning, S., 1996. GIS and Environmental Modeling: Progress and Research Issues. GIS World Books, Fort Collins, CO.

Green, M., 1990. Statistical methods for areal interpolation. In: Harts, J., Ottens, H.F.L., Scholten, H.J. (Eds.), EGIS '90: Proceedings of the First European Conference on Geographical Information Systems. EGIS Foundation, Utrecht, The Netherlands, pp. 392–399.

Jensen, J.J., Sathisan, S.K., 1996. GIS applications for linking travel demand modeling and air quality analysis. In: Geographic Information Systems for Transportation (GIS-T) Symposium. Proceedings of the 1996 Geographic Information Systems for Transportation (GIS-T) Symposium. American Association of State Highway and Transportation Officials, Washington, DC.

Longley, P., Batty, M., 1996. Spatial Analysis: Modeling in a GIS Environment. John Wiley and Sons Inc, New York, NY.

McCullagh, P., Nelder, J.A., 1989. Generalized Linear Models. Chapman and Hall, London.

Morris, R.E., Myers, T.C., 1990. User's Guide for the Urban Airshed Model. Volume I: User's Manual for UAM (CB-IV). Prepared for the U.S. EPA (EPA-450/4-90-007A) by Systems Applications, Inc., San Rafael, California.

Peuker, T.K., Fowler, R.J., Little, J.J., Mark, D.M., 1978. The triangulated irregular network. In Proceedings of the DTM Symposium, American Society of Photogrammetry-American Congress on Survey and Mapping.

Rao, S.T., 1987. Application of the Urban Airshed Model to the New

York Metropolitan Area. Prepared for the US EPA (EPA 450 4-87-011).

SAS, 1993. SAS/STAT Software: The GENMOD Procedure. Release 6.09, SAS Technical Report No. P-243. SAS Institute, Inc.

Scheffe, R.D., 1990. Urban Airshed Model Study of Five Cities (Summary Report). Prepared for the U.S. EPA (EPA-450/4-90-006A).

Shimp, D.R., Campbell, S.G., 1995. Using a Geographic Information System to Evaluate PM_{10} Area Source Emissions. California Air Resources Board, Sacramento, California.

Souleyrette, R.R., 1992. GIS for transportation and air quality analysis. In: Transportation Planning and Air Quality. American Society of Civil Engineers, New York.

Stata, 1998. Stata Reference Manual. Stata Inc, Santa Monica, California.

USEPA, 1975. Methodologies for Countywide Estimation of Coal, Gas, and Organic Solvent Consumption. EPA-450/3-75-086 (NTIS PB259909), U.S. Environmental Protection Agency, Research Triangle Park, North Carolina.

USEPA, 1995. National Conference on Environmental Problem-Solving with Geographic Information Systems, Cincinnati, Ohio. U.S. Environmental Protection Agency, Office of Research and Development.

USEPA, 1998. Better Assessment Science Integrating Point and Nonpoint Sources. EPA-823-B-98-006, U.S. Environmental Protection Agency, 401 M Street, SW, Washington, DC 20460.

Venables, W.N., Ripley, B.D., 1987. Modern Applied Statistics with S-Plus. Springer, New York.

Winkelmann, R., 1994. Count Data Models. Springer-Verlag, Berlin.