Journal of Environmental Sciences Vol.17, No.5, pp.861—865, 2005

ISSN 1001—0742
CN11—2629/X

Article ID: 1001-0742(2005)05-0861-05 **CLC number**: X21; X16 **Document code**: A

A model supported by GIS for locating and quantifying $PM_{2.5}$ emission originated from residential wood burning

ZHOU De-min[1,2,*], RADKE John[2], TIAN Yong-qian[3], XU Jian-chun[2], MU Lan[2]

(1. Northeast Institute of Geography and Agricultural Ecology, Chinese Academy Sciences, Changchun 130012, China. E-mail: zhoudemin@neigae.ac.cn; 2. Geographical Information Science Center, 102 Wheeler Hall, University of California, Berkeley, CA 94720, USA; 3. Center for Assessment and Monitoring of Forest and Environmental Resources, 151 Hilgard Hall, University of California, Berkeley, CA 94720, USA)

Abstract: A research method was presented for spatially quantifying and allocating the potential activity of a fine particle matter emission ($PM_{2.5}$), which originated from residential wood burning (RWB) in this study. Demographic, hypsographic, climatic and topographic data were compiled and processed within a geographic information system (GIS), and as independent variables put into a linear regression model for describing spatial distribution of the potential activity of residential wood burning as primary heating source. In order to improve the estimation, the classifications of urban, suburban and rural were redefined to meet the specifications of this application. Also, several definitions of forest accessibility were tested for estimation. The results suggested that the potential activity of RWB was mostly determined by elevation of a location, forest accessibility, urban/non-urban position, climatic conditions and several demographic variables. The linear regression model could explain approximately 86% of the variation of surveyed potential activity of RWB. The analysis results were validated by employing survey data collected mainly from a WebGIS based phone interview over the study area in central California. Based on lots free public GIS data, the model provided an easy and ideal tool for geographic researchers, environmental planners and administrators to understand where and how much $PM_{2.5}$ emission from RWB was contributed to air quality. With this knowledge they could identify regions of concern, and better plan mitigation strategies to improve air quality. Furthermore, it allows for future adjustment on some parameters as the spatial analysis method is implemented in the different regions or various eco-social models.

Keywords: residential wood burning(RWB); $PM_{2.5}$; demographical characteristies; Geographic information system(GIS); stepwise linear regression

Introduction

Wide scope of particulate emissions, from geologic particulates emissions, ammonia emissions, biogenic particulate emissions to smoke and combustion emissions were concerned by scientists presently because of more and more serious air pollution problems. Residential wood burning (RWB) contributes to the ambient concentration of fine particulate matter($PM_{2.5}$, the term used for a mixture of solid particles and liquid droplets found in the air, refers to particulate matter that is 2.5 μm or smaller in size) and volatile organic compounds (VOCs) (Michael, 2003; McDonald, 2000; Benner, 1995), which have been proven to have a major impact on human health (Browning, 1990). Although impacts of RWB on environmental air quality has drawn attention from scientists because a multitude of toxic substances and carcinogens were emitted into atmosphere due to rapid population growing in the past two decades (Purvis, 2000), air pollution generated from RWB remains inadequately addressed in most areas. It is urgent to have well defined information about where emission sources are located, so that the mitigation management against atmospheric degradation could be carried out efficiently (Diem, 2001).

Because of hardly defined environmental variables and complex residential behaviors in different regions, research of quantifying and locating on $PM_{2.5}$ from WRB is seldom tried. Some preliminary research mostly focused on emission data collection and estimation method improvement without qualified spatial variables involved. Current methods of estimating $PM_{2.5}$ from RWB are or the most part based on survey data (Houck, 2001). However, statistical method is usually costly and time-consuming, especially in large sampling areas, and the surveyed results can rapidly become outdated as demographics changed. Furthermore, the principles of emission can not be presented precisely or comprehensively if relative spatial variables are not involved, since the characteristics of RWB emission always highly correlated with geographic positions (Healey, 1990). There is an urgent need for a standard method of analysis for the amount and location of fuels being consumed by residential wood burning activities. Even qualifying variables of potential activity will provide improved emission estimates spatially. GIS based model of RWB emission sources (point, area, mobile) could include a spatial component. This approach provided tremendous benefits in making the data easier to use, update, display and share with people, and made it available for analysis of spatial characteristics of RWB emission variables and relationships between them while it was used for collecting, storing, processing, and visualizing the emissions input and output data (Patrick, 2001).

The activity of RWB was defined in this paper as the number of households with burning activity in a year. The method was linear regression for modeling the spatial distribution of potential activity of RWB as a function of demographic and physical environmental data, which were gathered from several existing databases and processed on the platform of ArcGIS 8.3. Some methods in the paper such as the GIS based clustering method for classifying urban-suburban-rural location, various definitions of forest accessibility and GIS based phone survey method could be useful in other research projects. The research results helped us to understand our air pollution activity more completely. This in turn led to develop more effective methods for reducing pollution and improving the air quality (Gene, 2000). For example, map-based analyses, combining

Foundation item: The research contract from California Air Resources Board (ARB), USA; the Talented Foundation from Northeast Institute of Geography and Agricultural Ecology, Chinese Academy of Sciences, China (No. C08Y17); * Corresponding author

demographic information such as income levels or urban/rural delineation to a GIS layer showing air toxic risks, provided a powerful analysis tool for exploring and helping to solve environmental justice issues or identifying where controls could provide the greatest health benefits to the most people.

1 Methods

1.1 Study site

Research objective in this study is to understand potential activity of RWB which occurred in the state of California. A sample site that includes 28 counties in the central California State(Fig.1) finally was chosen to describe the variety of California. It represented a good cross-section of California while it was characterized by multiple patterns of regional demographics, including the metropolitan area such as bay area and the capital of the state, Sacramento; it contained a suitable proportion of coasts and inlands, the metropolitan area in the western coastal area and small towns and countryside in the eastern inland, The rural area of the region was characterized by multiple land uses, with about 30% being forested areas or national parks that supply wood-fuel; and it also represented various climate regions known as the Sacramento Valley, the San Joaquin Valley and the San Francisco Valley.

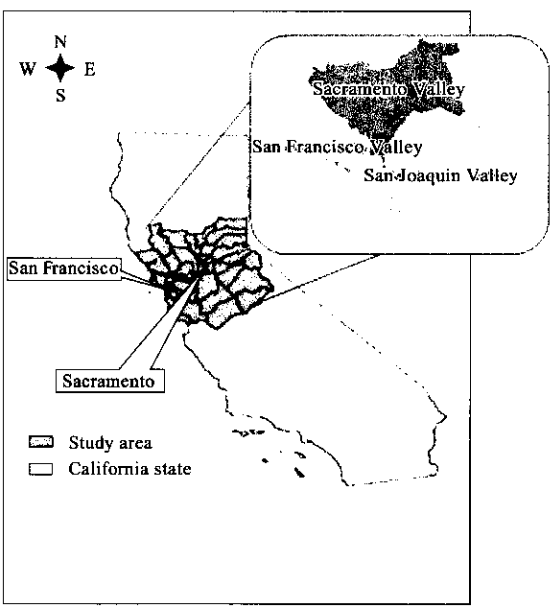

Fig.1 Study site in California

1.2 Data collection and compilation

Demographic factors were the prior choices for variables selection since our research of RWB was limited in "the residential". Demographic data for the State of California was obtained from the US 1990 Census. Characteristics of these demographics cover categories of population, housing, facilities, infrastructure, and socio-economic factors. Among the independent variables in the census, we observed evidence of RWB, our dependent variable. Although this variable is reported nationally as the number of households using coal or wood as their primary heating fuel, since no household has used coal as a source of fuel in California in the past 20 years, it represents residential wood burning. Due to the big range of census tract area in our study site, we normalized RWB and reported it as a density value for each census tract.

Whether RWB occurred or not mostly depended on the availability of alternative heating infrastructures such as gas and electricity. These utilities, for the most part, are heavily concentrated in urban regions where demand is high; and the distribution of other utilities, such as road networks, are also well developed in rural area. Only a two-level classification scheme of urban, rural was provided by US Census, so we could not observe the RWB activity in the suburban classification, which existed in the most study area. On the other hand, this scheme was based solely on population density(it might bias our analyzing on impacts from regional alternative heating infrastructures on RWB). So we redefined urban, suburban and rural areas by cluster analysis method (Wang, 1998) at census tract level based on 7 variables (population density, ratio of public water consuming, road density, ratio of bottle gas consuming, ratio of renters, ratio of farm income, median rent price).

Temperature and climatic conditions in general decided whether or not for heating. Climate data was obtained in the form of climatic zones provided by the California Energy Commission and represented as an average value over ten years. We adopt their concept of heating-degree-days(HDD, the number of degrees that the average daily temperature is below 65℃, with which heating is likely to occur). These HDD are much more likely to occur at higher elevations where recorded temperatures are consistently lower in value. The elevation for the study area was obtained from the USGS as a digital elevation model(DEM) with a scale of 1:250000(a resolution of approximately 83 m × 83 m), which was adequate for our regional scope.

RWB will be regularly occurred in regions where fuel wood is accessible. The majority of wood-fuel sale originates from the sites maintained by the US Forest Service, which contributes approximately 70% of RWB energy. We sought to discover how RWB potential activity was related to accessibility to forest fuel-wood supply. We selected the attribute "Forested Upland" to represent forestland cover, it came from the USGS National Land Cover Data (NLCD). Accessibility to forest wood-fuel was calculated and compared using the radius distance from residents to the nearest forest boundary, and the percentage of forested areas within a specified radius distance(about 30 mile radius in our study).

1.3 Stepwise linear regression modeling

To identify important independent variables that could be used to predict residential wood burning, a stepwise linear regression analysis was conducted. The candidate independent variables for the statistical analysis included demographic characteristics, elevation, climatic zones, location of urban, suburban and rural areas, and generated information on accessibility to forest fuel-wood. Both forest accessibility and radius distance to the nearest forest edge were included as independent variables. The dependent variable, the number of households using wood as primary heating source, was obtained from the 1990 census data. Without considering correlation effects, Equation (1) below explained approximately 86% of the variation with the top seven variables selected from all candidate variables by using stepwise linear regression method:

$$AP = 3069 + 2.342x_1 + 412.75x_2 + 102.96x_3 + 30x_4 - 45x_5 + 126.6x_6 - 23.5x_7, \quad (1)$$

where x_1 represents elevation; x_2 is the locations of urban, or suburban, or rural; x_3 is the ratio of retirement residents; x_4 is the forest accessibility; x_5 is the ratio of houses occupied by its owner; x_6 is the ratio of households living on farming incomes; x_7 is the percentage of households in which the age of the head of the family is greater than 55.

2 Results

The model demonstrated that that potential activity of RWB was predictable. Among all variables, elevation was the most influential since cold and windy climates at high elevations generate high demand for heating. This quantitative evidence is consistent with phenomena described by Murphy *et al*. (Murphy, 1984). In addition, most of the forests (potential sources of fuel-wood) are located in the highly elevated mountainous regions. Casual observation of wood-fuel sales reveals that supermarkets in the Lake Tahoe region sales more fuel-wood than that in San Francisco Bay Area counterparts, because Lake Tahoe is located on the border between California and Nevada and its average elevation is 6229 feet above sea level. These data approved our study result.

Next to elevation, forest accessibility explained approximately 56% variation of residential wood burning potential activity. However, a high correlation between elevation and forest accessibility caused this variable to be ranked 4th in the stepwise analysis. The higher number of households using wood burning that related to forest accessibility certainly also related to elevation and urban/rural land use types. Our linear regression analysis showed that elevation had more influences than forest accessibility, although both variables were good predictors of RWB potential activity. Sexton *et al*. (Sexton, 1985) reported that the mean monthly temperature was important factor influencing wood burning activity of households. The high correlation among variables of elevation, HDD and forest accessibility rendered HDD insignificant when elevation and forest accessibility variables were included.

Several demographic characteristics were directly related to RWB practices: proportions of retired income (RETI_INC), ratio of occupied houses, farm income in households (FARM_INC), age groups of residents (AGE > 55 or 34 < AGE < 54), and owner occupied units (OWNR_OCC). The contributions to the predictive model using the selected demographic variables in each scenario were consistent. Retired people require more economic heat energy. Farm households have better accessibility to wood fuels. Occupied units are an obvious factor since likely only occupants need heat energy. Residents whose ages range between 35 or 54 are more likely to have small children, have pressures of income, and are more likely to look for cost effective fuels.

3 Validations

The stepwise linear regression method was a valuable first step in identifying independent variables that correlated with potential activity of RWB. These identified important variables were satisfactory to expert conceptual assessment. In order to calibrate and validate the model and collect updated information on the distribution and variation of RWB activity, a survey was designed and given in person or over the telephone to households in the study area.

The survey information mainly included location, duration of average burning event for both weekdays and weekends, estimated number of annual burning events, and months that burning occurred. Demographic information required by survey included type and age of dwelling, availability of wood burning fire appliances, tenure status, age of the oldest family member, and education level. A stratified random sampling method was employed for generating survey sample points in our study area because of the large variation in household density. The stratification was based on areas classified as urban, suburban and rural to ensure an even distribution and a sufficient number of points for each category. In case of potential anomalies, households close to the sampling points were visited or telephoned to until three successful samples were completed. The survey started from traditional method, face to face interview. A WebGIS based phone number collection method integrated into a stratified random sampling protocol was applied to survey strategy, which achieved big progress. WebGIS based residential phone directory supported by Pacific Bell Phone Company made it easy to generate phone number of each random survey point over the entire study region. Over 15000 survey interviews, including phone calls and face-to-face visits, resulted in 731 successful acquired samples; most of them produced by phone survey. The success rate from face-to-face visits was higher than that from phone interviews, but with much more time and higher cost. Lots survey points lost because the survey principle of spatial random made it possible that some points were not in residential area, especially in suburb and rural area.

The 731 successfully completed samples were used for validating the linear model (Eq. (1)) of estimating potential activity of RWB. The validation was conducted for four spatial circumstances at study site: (1) urban and non-urban locations; (2) individual counties; (3) climatic zones; (4) terrain elevation classes.

On the view of the Model performance in urban and non-urban classifications for three regions in our study (Table 1), the model provided excellent estimates of surveyed percentage of households burning wood as a primary heating source in all three regions ($R^2 > 0.97$). However, the model overestimated for the non-urban areas about 2 times and underestimated the urban areas. Although stratified random sampling methods were applied to ensure the even distribution of the sampling points, survey response rate in forested non-urban area was very low due to the poor quality phone directory and losses of survey points. These low-response forested areas usually have a high percentage of households burning wood as primary heating source. Therefore, the surveyed data underrepresented the true conditions in non-urban locations.

Sixteen counties where we have sufficient survey points are included to validate the model prediction for potential residential wood burning (Table 2). The model explained about 68% of the variations of the survey data. Model estimation matched the survey data very well in almost all counties. We observed that elevation and forest accessibility could obviously affect model prediction. For example, in Nevada County, the high burning percentage in the survey matched model estimation because that Nevada County is a heavily forested area with high elevation; however, in Calaveras County, estimated values were much higher than those from the surveyed data. This unmatched situation caused by those successfully completed samples clustered in low elevation and non-forested areas.

Table 1 Percentage of households burning wood as primary heating source

Location	Rural/suburb		Urban		Urban and non-urban	
	Survey	Model	Survey	Model	Survey	Model
Sacramento Basin	12.94%	24.34%	5.45%	1.96%	9.56%	14.14%
San Francisco Bay	5.00%	10.06%	4.57%	1.17%	5.16%	2.72%
San Joaquin	10.93%	21.90%	6.67%	2.48%	10.09%	17.83%
Correlation	0.996		0.978		0.990	

Table 2 Percentage of wood burning households from survey and model prediction at county level

County	Survey, %	Model, %	County	Survey, %	Model, %
Alameda	7	1.45	Contra Costa	3	2.83
Merced	19	11.39	Marin	0	6.45
Nevada	29	37.79	San Mateo	4	6.11
Placer	20	24.75	Santa Clara	4	0.96
Sacramento	5	2.02	Solano	8	10.37
San Joaquin	12	9.04	Sonoma	8	10.42
Stanislaus	6	10.06	Sutter	5	8.61
Calaveras	17	39.44	Yolo	8	6.39

Six different climatic zones existed within the study area and their mean heating-degree-days (HDD) varied from 2643 to 5532(Table 3). One of these zones is an unusual region covered mainly by high elevation forest and has average HDD of 5532. Validation in the unusual climatic zone was excluded because of insufficient sample points in this region. The model predicated and surveyed ratios of wood burning households for each climatic zone had a correlation coefficient of 0.786. Two overestimations were in climatic zones in northern California and two underestimations were in coastal regions around central California.

Table 3 Model predicted and surveyed ratios of wood burning households in individual climatic zones

Climate Zone	Model, %	Survey, %	HDD
2	10.28	2.78	3026
3	2.85	4.17	2840
4	0.96	3.92	2643
11	23.32	16.67	2847
12	10.73	8.24	2697

The study site could be naturally classified into five categories of DEM on the classification rules(Table 4). The correlation between surveyed and predicted ratios of wood burning households for each DEM class was 0.77. The model did not seem to perform better with DEM variation than in scenarios of climatic zones and counties. This phenomenon appeared inconsistent with the regression result, which suggested that elevation was the most important factor in predicting potential activity of RWB. Actually, the two results described different things, because this analysis depended on the rules of classifying the DEM. This analysis was not intended to find the best classification rule. Instead, it was to use a classification case to examine the model performance against the survey result. This validation showed that the model performed satisfactorily with the variation of the DEM.

Table 4 Model predicted and surveyed the ratios of wood burning households for five DEM categories

DEM Class	Class rule	Model, %	Survey, %
1	0—77	4.22	31.88
2	77—244	11.81	34.26
3	244—488	30.81	28.13
4	488—798	38.22	36.11
5	798—1694	59.48	75.00

4 Conclusions and discussion

This study discovered that potential activity of residential wood burning (RWB) was a function of spatial and demographic characteristics. The most important spatial characteristics were elevation, accessibility to forest areas, urban/non-urban classification, and climate conditions. Elevation was the best predictor, with forest accessibility as a strong second place predictor. Variation of elevation implied change of local climate that increased or decreased demand for heat in a household. The forest accessibility affected potential activity of RWB because firewood was a cheap heating resource and convenient access to the resources could change heating habits. The analysis results also indicated that demographic characteristics were complementary to spatial characteristics in predicting RWB potential activity. The differences of potential activity between urban and rural areas were significant, Roger *et al*. (Roger, 2004) reported the similar research result.

The methods for extracting spatial and demographic variables were critical for allocating and qualifying $PM_{2.5}$ emission from RWB. From the number of primary wood burning households in a year (the conception of activity of RWB in this paper), and other two variables of fuel-wood consumption and combustion efficiency of fire appliances, $PM_{2.5}$ emission in a specific region could be estimated from some general formula. Actually, the study on the later two variables have been well done(Cass, 2000), however, our study is the first effort.

The suburban was observed as an independent pattern for its potential activity of RWB in our research. The modeling results demonstrated that the potential activity of RWB in suburban area matched that in rural area rather than urban area, although residents in suburban area, living at the edge of urban, kept close relationship with the urban. This phenomenon probably could be explained by that residents in suburban area usually love the rural life style although most of them get their jobs in cities (Healey, 1990). "Personal style" or "residential habit" sometimes determines various aspect of activity of RWB such as fire appliances, the type of fuel-wood, the time to start burning and the motive to burn for aesthetic or heating. From these points, we recognized the complexity of "the residential" (Kimmo, 2001), the core of which tends to be nonlinear and is not efficient to apply with linear method to describe. Although linear analysis is an important step to identify significant factors affecting human use of wood as a heat source, and the model is sufficient for qualifying potential activity of RWB. Some incompatible results still convinced us that a further study using a non-linear modeling approach might be appropriate for describing the mechanics of the causal relationship in this human social ecological system (Eduardas, 1998). The time scale is the other problem remained to be further study. Since our model is based on the annual time scale, it obviously could not demonstrate seasonal change of $PM_{2.5}$ emission, which is highly concerned by scientists and administrators, monthly or daily scale models would be our next goal (Roger, 2004; Robert, 2003; Radke, 2002).

Acknowledgement: This research was conducted during first author's academic visit in GISC, University of California Berkeley. We are grateful to Don Rake, Patrick Gaffney and Klaus Scott for providing the research group with valuable suggestions. Houck helped design the initial survey questions and commented on the later drafts.

References:

Benner B A, Wise S A, Currie L A *et al.*, 1995. Distinguishing the contributions of residential wood combustion acid mobile source emissions using relative concentrations of dimethylphenanthrene isomers[J]. Environ Sci and Technol, 29(9): 2382—2389.

Browning K G, Koenig J Q, Checkoway H *et al.*, 1990. A questionnaire study of respiratory health in areas of high and low ambient wood smoke pollution[J]. Pediatric Asthma Allergy and Immunology, 4: 183—191.

Cass G R, Fine P M, 2000. The contribution of biomass combustion to ambient fine particle concentrations in the United States [EB]. EPA-R826233. http://es.epa.gov/ncer/progress/ grants/97/airqual/cass00.html.

Diem J E, Comrie A C, 2001. Allocating anthropogenic pollutant emissions over space: application to ozone pollution management [J]. Journal of Environmental Management, 63(4): 425—447.

Eduardas K, Lee S, Stephen M, 1998, The residential space heating problem in Lithuania[J]. Energy Policy, 26(11): 831—858.

Gene T, 2000. An overview of $PM_{2.5}$ sources and control strategies[J]. Fuel Processing Technology, (65/66): 379—392.

Healey M, Ilbery B, 1990. Location and change: Perspective on economic geography[M]. Oxford, Reading: Oxford University Press.

Houck J E, Mangino J E, Brooks G *et al.*, 2001. Recommended procedure for compiling emission inventory national, regional and county level activity data for the residential wood combustion source category[C]. In: Proceedings of the US environmental protection agency emission inventory conference, Denver, CO.

Kimmo J K, Otto H, Tuulia R *et al.*, 2001. Behavioral and environmental determinants of personal exposures to $PM_{2.5}$ in EXPOLIS—Helsinki, Finland [J]. Atmospheric Environment, 35(14): 2473—2481.

McDonald J D, Zielinska B, Fujita E M *et al.*, 2000. Fine particle and gaseous emission rates from residential wood combustion[J]. Envirn Sci Technol, 34: 2080—2091.

Michael D H, Dean S, John K *et al.*, 2003. Polycyclic aromatic hydrocarbon size distributions in aerosols from appliances of residential wood combustion as determined by direct thermal desorption—GC/MS[J]. Journal of Aerosol Science, 34(8): 1061—1084.

Murphy D J, Buchan R M, Fox D, 1984. Ambient total suspended particulate matter and benzo(a) pyrene concentrations from residential wood combustion in a mountain resort community[J]. American Industrial Hygiene Association Journal, 45: 431—435.

Patrick G, 2001. A comprehensive research plan for developing $PM_{2.5}$ emission inventories[EB]. 10th Annual Emission Inventory Conference, http://www.epa.gov/ttn/chief/conference/ ei10/pm/gaffneypres.pdf

Purvis C R, McCrillis R C, 2000. Fine particulate matter(PM) and organic speciation of fireplace emissions[M]. Environ Sci and Technol, 34(9): 1653—1658.

Radke J D, Gong P, Tian Y Q, 2002. Model development for spatial allocation of $PM_{2.5}$ emissions including residential wood combustion [R]. State of California Air Resources Board. Technical Report, #99—716.

Robert G, Brigitte B, 2003. Characterising seasonal variations and spatial distribution of ambient PM_{10} and $PM_{2.5}$ concentrations based on long-term Swiss monitoring data[J]. Atmospheric Environment, 37(19): 2571—2580.

Roger L T, William J P, Myra L V *et al.*, 2004, Regional composition of $PM_{2.5}$ aerosols measured at urban, rural and "background" sites in the Tennessee valley[J]. Atmospheric Environment, 38(20): 3143—3153

Sexton K, Liu K S, Hayward S B *et al.*, 1985. Characterization and source apportionment of wintertime aerosol in wood-burning community [J]. Atmospheric Environment, 19(8): 1225—1236.

Wang Q, Hu S, 1998. The model of clustering for classification[M]. In: Modeling and management of geographic information science applied by regional planning. Beijing: Yuhang Press. 55—57.

(Received for review January 3, 2005. Accepted February 6, 2005)