14th International Conference on
Harmonisation within Atmospheric Dispersion
Modelling for Regulatory Purposes
Kos Island, Greece, 2-6 October 2011

MAPPING TRAFFIC ATMOSPHERIC EMISSIONS FOR EPIDEMIOLOGICAL STUDIES USING ATMOSPHERIC DISPERSION MODELS AND GEOSTATISTICAL METHODS: A CASE STUDY

Sergio Teggi[1], Grazia Ghermandi[1], Sara Fabbi[1], Carlotta Malagoli[2], Marco Vinceti[2] Luisa Guerra[3], Antonella Sterni[3] , Giuseppe Maffeis[4]

[1]Department of Mechanical and Civil Engineering, University of Modena and Reggio Emilia, Modena, Italy, e-mail: sergio.teggi@unimore.it
[2]Research Center for Environmental, Genetic and Nutritional Epidemiology, Dep. of Public Health Sciences, University of Modena and Reggio Emilia, Modena, Italy
[3]ARPA Emilia Romagna, Sezione di Modena, Modena, Italy
[4]TerrAria s.r.l., Milano, Italy

Abstract: In some cases, epidemiological studies require the air pollutant concentrations at the exposure points. In these cases air dispersion models represent a very important tool. When additional points of exposure are inserted or when some exposure points must be relocated, spatial interpolators can be used in place of new runs of the air dispersion model. In this work the uncertainties and the problematic related to spatial interpolation methods are inspected. The case studied is based on an epidemiological study aimed to study the risk of childhood leukemia associated with benzene exposure due to traffic emissions. The concentration values of benzene computed by the atmospheric dispersion model ADMS are taken as reference and compared with the concentration values computed using several interpolation methods and additional data sets of concentrations computed by ADMS in the same area. The comparison is done following two approaches: the summary statistics of the differences and the correctness of the assignment of the exposure points to the concentration categories used in the epidemiological study. These comparisons show that the values computed by the interpolators are very problematic: important differences and categories assignment and categories uncertainties were found.

INTRODUCTION

Air dispersion models estimate pollutants concentration at specific locations (receptors) or at locations distributed on spatial grids. Often, some exposure points have to be relocated or additional exposure points are requested: in these cases air dispersion model must run again, or the concentration on the new locations can be computed by spatial interpolation of the values already available. Therefore, interpolation methods can play an important role in exposure calculation and the reliability assessment of the interpolated values is crucial. This study (Malagoli et al., 2011, Vinceti et al., 2011) is aimed to evaluate the risk of childhood leukemia associated with benzene exposure due to traffic emissions in the Modena province (Italy). The concentration of benzene at the exposure points are computed using the ADMS (McHugh et al., 1997) air dispersion model. The study area is located in the northern part of the Modena Province (Po Valley, Italy), has an extension of 55 km x 60 km and comprises the most populated zones of the Province. We computed new concentration values at the exposure points using several interpolation methods and the ADMS values computed at different locations.

DATA SET

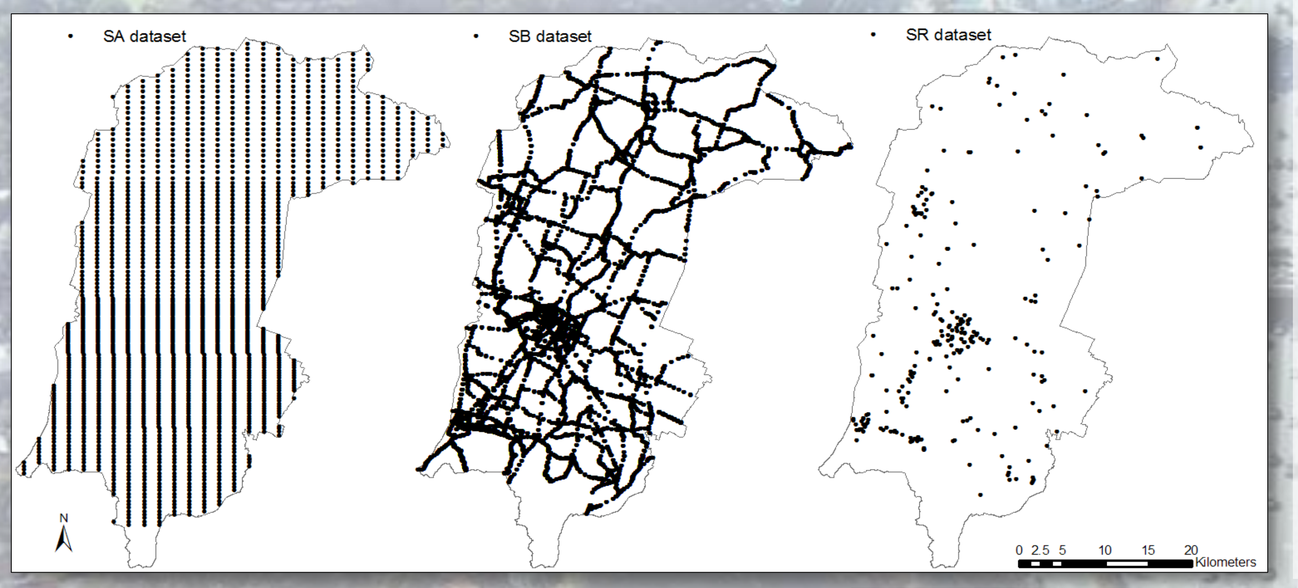
a) 2077 points located at the regular grid (SA, Figure-left);

b) 19777 points located at the intelligent grid (SB, Figure-center);

c) 4220 points obtained by spatial aggregation of SB points (SC)*;

d) 240 validation points located at the exposure points (or receptors, SR, Figure -right).

* The aggregation is done using a blocking method: the domain is dived into 150 m x 150 m cells (blocks) and then all the points contained in a cell are replaced by a point located at mean position and with the mean concentration value of the original points.



INTERPOLATION AND COMPARYSON METHODOLGY

Interpolation methods to compute the concentration at the exposure points (SR dataset), (Isaaks and Srivastava, 1989; O'Sullivan and Unwin, 2003):

- Voronoj polygons (**VO**);
- Inverse Distance method (**ID**);
- local Linear Interpolation (**LI**, first order);
- S-Pline (**SP**);
- Kriging (**KR**);
- Co-Kriging (**CK**).

The interpolation methods are applied to 3 data sets: **SA** data set, **S1** data set (SA+SB) and **S2** data set (SA+SC). CK was applied in 2 different configurations: in both the cases the SA data set was used as the principal variable, while the correlated variable was set to SB in the first case (**CK1**) and to SC in the second case (**CK2**).
It computed separate statistics for each of the concentration categories (I_i) used in the reference epidemiological study: $\mathbf{I_0}$ = [0, 0.1] μg/m^3, $\mathbf{I_1}$ =]0.1, 0.5] μg/m^3, $\mathbf{I_2}$ =]0.5, 1.0] μg/m^3 and $\mathbf{I_3}$ =]1.0, ∞] μg/m^3. The first category has been added with respect to those used in the reference study in order to consider a null (insignificant) concentration level. For each of these categories several statistical parameters have been computed using as reference the ADMS data set :

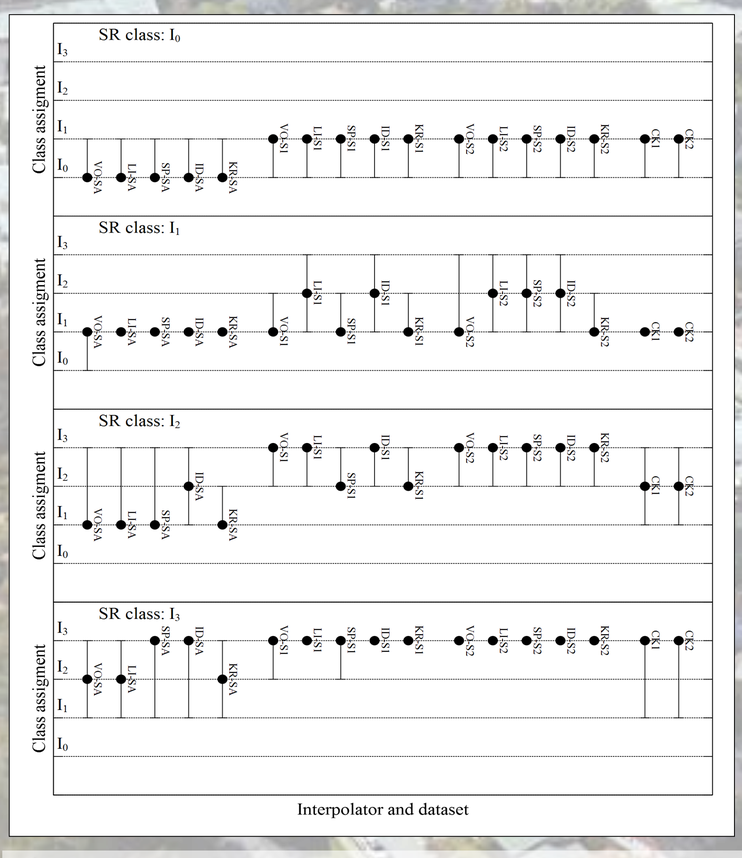
1. $\mathbf{N_{ci}}$, the number of exposure points correctly assigned by a given interpolator to the category I_i;
2. $\mathbf{\Delta_i}$ and $\mathbf{RMSD_i}$, the bias and the root mean square difference between the concentrations interpolated and those computed by ADMS of the exposure points belonging to the ADMS data set I_i;
3. **Median**, 0.16 and 0.84 quantiles (these quantiles include the 68% of the cases) of category number, determined by the interpolated value, of the exposure points belonging to the ADMS I_i data set.

Category	I_0			I_1			I_2			I_3		
$Range_{SR}$	0.06 ± 0.02			0.24 ± 0.11			0.68 ± 0.13			1.79 ± 0.88		
N_{SR}	52			127			41			20		
	N_{C0}	Δ_0	$RMSD_0$	N_{C1}	Δ_1	$RMSD_1$	N_{C2}	Δ_2	$RMSD_2$	N_{C3}	Δ_3	$RMSD_3$
VO-SA	41	0.03	0.14	92	0.01	0.21	7	-0.03	0.63	7	-0.66	1.5
LI-SA	33	0.03	0.09	106	-0.01	0.12	11	-0.06	0.49	9	-0.84	0.91
SP-SA	33	0.03	0.06	107	-0.01	0.12	11	-0.05	0.5	10	-0.74	0.98
ID-SA	33	0.03	0.06	106	0	0.12	12	-0.05	0.5	10	-0.76	0.97
KR-SA	33	0.02	0.04	112	-0.02	0.11	14	-0.11	0.48	6	-0.83	0.93
VO-S1	26	0.06	0.1	79	0.25	0.37	17	0.51	0.62	16	0.08	0.82
LI-S1	21	0.12	0.21	52	0.42	0.39	8	0.77	0.67	18	0.37	0.9
SP-S1	17	0.04	0.06	91	0.15	0.19	30	0.21	0.41	14	-0.25	0.71
ID-S1	14	0.15	0.15	37	0.48	0.37	9	0.81	0.74	18	0.29	0.94
KR-S1	14	0.05	0.06	79	0.23	0.24	23	0.39	0.53	17	0.02	0.81
VO-S2	24	0.09	0.15	69	0.32	0.4	12	0.75	0.89	18	0.28	0.86
LI-S2	18	0.11	0.15	48	0.4	0.32	8	0.68	0.63	18	0.16	0.9
SP-S2	16	0.1	0.12	56	0.37	0.38	12	0.66	0.66	19	0.22	0.84
ID-S2	14	0.14	0.15	50	0.4	0.33	9	0.7	0.61	18	0.22	0.86
KR-S2	14	0.09	0.1	65	0.34	0.38	13	0.67	0.8	19	0.24	0.98
CK1	14	0.02	0.06	115	0.01	0.12	16	0.01	0.54	11	-0.56	0.83
CK2	14	0.03	0.06	112	0.04	0.13	13	0.09	0.64	12	-0.59	0.93

RESULTS

The exposure points with "not significant" concentration levels, $\mathbf{I_0}$, are better approximated using the SA dataset only. In this case more than 60% of the exposure points belonging to the I_0 category are correctly assigned (78% in the VO case).
For the estimates of exposure points belonging to the $\mathbf{I_1}$ category, the percentage is 88% for the KR interpolator and 84% for the others but VO (72%) using the SA dataset. The same performances are achieved using the CK method.
For the third category, $\mathbf{I_2}$, the % is lower than 42% in most of the cases, the concentrations obtained with the S1 and S2 datasets are often overestimated. The only two exceptions are the SP-S1 (73%) and the KR-S1 case (56%).
In the last category, $\mathbf{I_3}$, the better performance is obtained using the S1 and the S2 datasets (excluding the VO-S1 and the SP-S1 cases). The % range from 85% to 95%.
The figure shows the plots of medians and of 0.16 and 0.84 quantiles (error bars, assignment error) of category assignments.

The exposure misclassifications arising from the different interpolation methods examined in the present study would have marked effects on computation of health risks attributable to the benzene emissions. In addition, since the misclassification of exposure occurring with the different methods would be differential, i.e. not characterized by a uniform pattern across the different strata I_i, these effects would not be a simple reduction in amount and statistical stability of the risk estimates, as expected in case of non-differential exposure misclassification, but the induction of severe bias in estimates computed for specific strata.



CONCLUISIONS

The use of atmospheric dispersion modelling is an important tool for the calculation of concentration levels of pollutants at the exposure points required by epidemiological studies. When additional exposure points are required or when the exposure points have to be relocated to approaches are commonly used: to make new runs of the model or to use spatial interpolators. The numerical comparisons of the two set of data showed substantial differences (bias and root mean square differences). Using the interpolated values, the assignment to the exposure categories, utilized in the epidemiological study, showed important discrepancies and uncertainties. These considerations can be carried out for all the interpolation methods used (except for the Voronoj method in pejorative sense) and for all the data sets used. In conclusion the use of interpolators must be done with extreme caution in epidemiological studies.

Reference

Isaaks, E. H. and Srivastava, R., M., 1989, An Introduction to Applied Geostatistics, Oxford University Press

Malagoli, C., Cherubini, A., Maffeis, G., Sterni, A., Guerra, L., Fabbi, S., Teggi, S., Ferretti, E. and Vinceti, M., 2011: Stima della esposizione a benzene da traffico veicolare nelle provincie di Modena e Reggio Emilia, XII Conferenza Nazionale di Sanità Pubblica, Rome, 12-15 October.

McHugh, C., Carruthers, D. and Edmunds, H., 1997, ADMS-Urban: an air quality management system for traffic, domestic and industrial pollution, International Journal of Environment and Pollution, 8, 666-674

O'Sullivan, D. and Unwin, D. J., 2003, Geographic Information Analysis, John Wiley & Sons

Vinceti, M., Malagoli, C., Sterni, A., Guerra, L., Cherubini, A., Maffeis, G., Fabbi, S., Teggi, S., Ferretti, E., De Girolamo, G., Palazzi , G. and Paolucci, P., 2011: Esposizione a benzene da traffico e leucemia infantile: influenza delle modellistiche di dispersione atmosferica sulla valutazione del rischio, XII Conferenza Nazionale di Sanità Pubblica, Rome, 12-15 October.