

Nitrogen dioxide and carbon monoxide estimation in the metropolitan area of Catania

Letizia La Tona¹, Romana Gargano¹, Luca S. Scarlata²

¹ Department of Economical, Financial, Social, Environmental, Statistical and Territorial Sciences, University of Messina

² Faculty of Statistics, University of Messina
e-mail: rgargano@unime.it

Abstract: The aim of this paper is to study the nitrogen dioxide and carbon monoxide concentrations in Catania from 2005 to 2007. A two stage approach is used to estimated its behaviour. In the first stage we estimated site-specific trends in data adjusted for the effects of season and meteorology and in the second stage, we estimated metropolitan area trend.

Keywords: Generalized additive model, air pollution, meteorological data.

1. Introduction

Air quality is an issue of primary importance for human health. Numerous epidemiological studies have assessed the health effects of air pollution and various public authorities have regulated air quality monitoring and settled standards. Hence, local public institutions have been lead to invest in networks to assess the concentration levels of various pollutants around the country. This paper describes a two-stage modelling approach to estimate trends in concentrations of nitrogen dioxide (NO₂) and carbon monoxide (CO) measured in Catania (Sicily, Italy). The aim is to identify the trends of NO₂ and CO daily concentrations adjusted for the effects of meteorological covariates in an area where there are not similar studies. In literature this kind of analysis is faced with space–time models. In our study we propose to use Generalized Additive Models, considering the shortness of time series, the small number of sites and that the temporal dimension is not of primary interest. These models identify and characterise non liner regression effects. The paper is organized as follow: the Section 2 contains a description of the data, the Section 3 illustrates the method used, the Section 4 shows main results and conclusions.

2. Study area

Catania is an important attraction center of eastern Sicily and it is a daily destination for many users. This area has particular features of conurbation with the hinterland and consequently the vehicle traffic is very high; in fact the resident population is 298,957 but the metropolitan area includes 752,895 inhabitants. Moreover the number of users is

double then the residents one and the ratio residents/vehicles is 695/1000 per inhabitants, more than two times as many cities as London or Tokyo.

The air pollutants analysed are NO₂ and CO recorded respectively at 12 and 14 monitoring sites in the municipal area, all of them are equipped with analogous instruments. These data were observed from 1 January 2005 to 31 December 2007. Data of the pollutants were furnished by XIII Department (Ecologia, Ambiente e Nettezza Urbana) of Catania Municipality. The observed meteorological variables are: air temperature (Temp), humidity (H), wind speed (WS), atmospheric pressure (AP), solar radiation (SR) and rain (Ra). In Catania there are only three weather monitoring stations, we use a mean of meteorological variables as indicator because the values are quite similar ($p=0.561$ of the differences between sites). All observations are available on an hourly basis, but in this work a daily mean value was found to be more adequate.

3. Model for municipal trend

Preliminary we describe spatially the air quality in Catania metropolitan area. To estimate the pollutants values in not sampled sites we use the local forecasts by means of inverse distance method. It gives greater weight to sample points that are close to the site of interest. We assume $p=2$.

Concentrations of nitrogen dioxide and carbon monoxide observed in Autumn-Winter and Spring-Summer are presented as contour plots extended over the areas of concern (Figures 1 and 2).

In order to estimate the form of relationship between the logarithm of daily NO₂ or CO concentrations to prevailing meteorological condition and time, we use a generalized additive model (Hastie and Tibshirani, 1990).

Figure 1: Contour plots for NO₂ Autumn - Winter (1a) and Spring - Summer (1b)

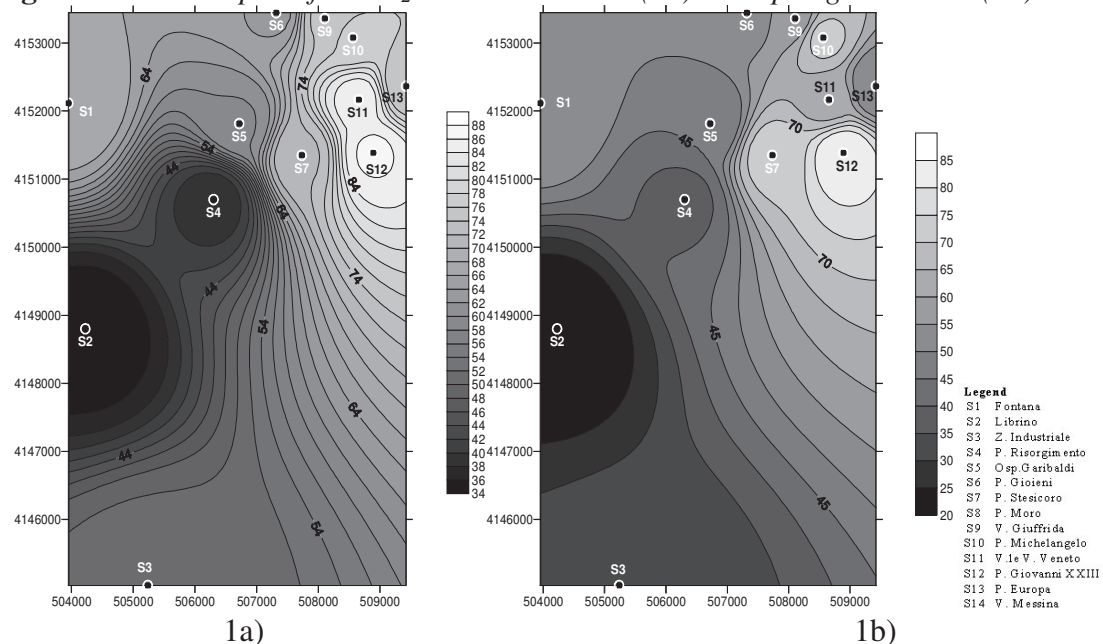
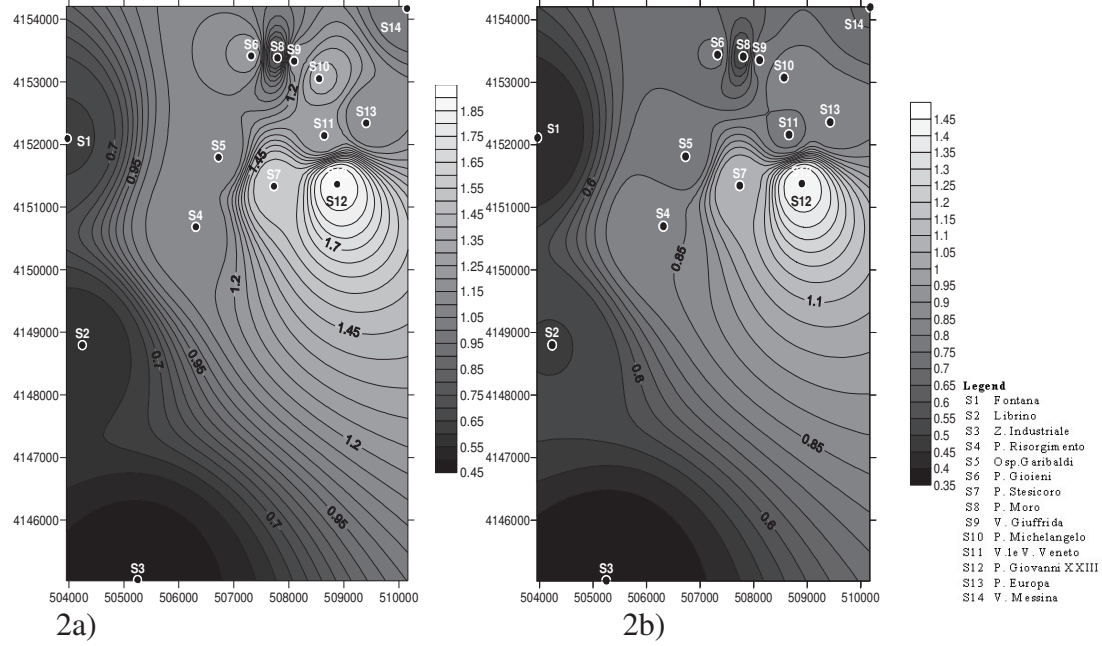


Figure 2: Contour plots for CO Autumn - Winter (1a) and Spring - Summer (1b)



According to Li et al. (2007) and Holland al. (2000), we use the natural logarithmic transformation in order to obtain a symmetric distribution and residuals in closer agreement with the distributional properties of the errors in model. Generalized additive models have been used as a standard analytic tool in time-series studies of air pollution and health during the last decade. A major statistical concern came into view recently about the appropriateness of the use of GAMs in the presence of concurvity, which is likely to be present in the data of all air pollution studies. These models can be virtually used in any setting in which linear models are used. They provide a flexible alternative to say fitting polynomials for investigating and identifying possible nonlinear relationships between a response variable and explanatory variables. Local smoothers, including smoothing splines, may be included in the regression function but estimation is not by maximum likelihood. Rather, a penalty term, which reflects the degree of smoothness in the regression, is added to the log-likelihood and this sum of terms is maximized.

In the first stage we estimate site-specific trends in data, the model is:

$$\log(Y_{ijl}) = \mu_l + s_{1l}(WS_{ijl}) + s_{2l}(Temp_{ijl}) + s_{3l}(H_{ijl}) + s_{4l}(Ra_{ijl}) + s_{5l}(AP_{ijl}) + s_{6l}(SR_{ijl}) + s_{7l}(Lag_{ijl}) + s_{8l}(Diff_{ijl}) + \varepsilon_{ijl} \quad (1)$$

where Y_{ijl} is the measured pollutant concentrations of the i^{th} day within the j^{th} month at the l^{th} monitoring station, $s_{kl}(\cdot)$, $k=1, \dots, 8$ are smooth functions of corresponding covariates and $\varepsilon_{ijl} \sim N(0; \sigma^2_{ijl})$ are independent random variables. We used the cubic splines as smooth function and, to reduce the strong correlation of residuals, we used the series of lag 1 (“Lag”) and the third differences (“Diff”) of the pollutants. For each station the model was fitted by a local scoring algorithm. The inclusion or exclusion of each variable and the smoothing method were been determined via a stepwise regression.

Next step, in order to estimate the whole municipal trend, we determine an area indicator based on the average of daily measurements of all stations. The model is:

$$\log(Y_{ijl}) = \mu_l + s_{1\mu_l}(WS_{ij\mu_l}) + s_{2\mu_l}(Temp_{ij\mu_l}) + s_{3\mu_l}(H_{ij\mu_l}) + s_{4\mu_l}(Ra_{ij\mu_l}) + s_{5\mu_l}(AP_{ij\mu_l}) + s_{6\mu_l}(SR_{ij\mu_l}) + s_{7\mu_l}(Lag_{ij\mu_l}) + s_{8\mu_l}(Diff_{ij\mu_l}) + \varepsilon_{ij\mu_l} \quad (2)$$

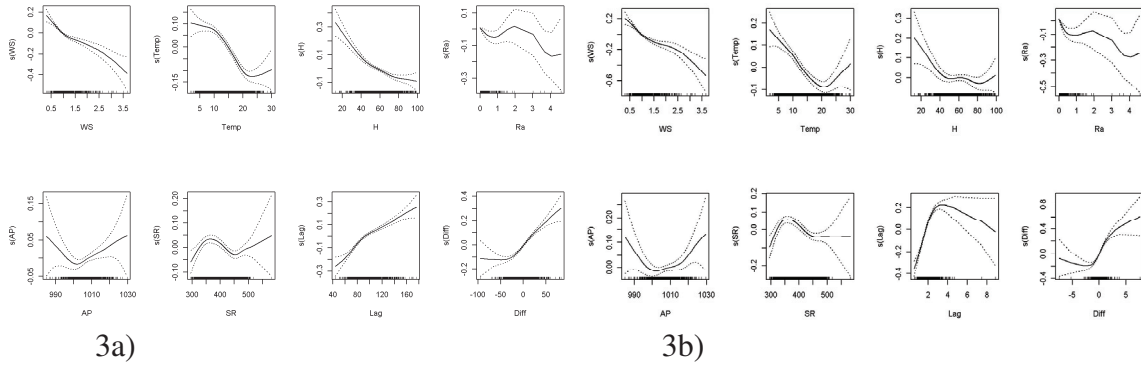
where μ_l is the concentration pollutant average.

4. Results and Conclusions

Contour plots show that the highest concentrations of NO₂ and CO are registered in S12, where some of the most important and busy roads of the city converge. The lowest values are in S2 and S4 stations, for NO₂, and in S1, S2, S3 and S8, for CO. All these monitoring stations are located out downtown.

Temperature, relative humidity and wind speed variables are important factors for all NO₂ GAM models. The relationships between CO and weather variables are less clear, in general, only wind speed and the solar radiation are important. Performed models show a sufficient goodness of fit in all monitoring stations. The GAM models of the area indicator confirm the results obtained in the first model. The form of the smoothed curves for each variable and confidence bands for the curves are shown in figure 3 (3a for NO₂ and 3b for CO).

Figure 3: Smoothed curves for NO₂ (3a), CO (3b) predictors and confidence bands



The negative relationships between the NO₂ and CO concentrations and wind speed suggests that independently of season the windy conditions decrease the concentrations. Presumably, the reason is that the increased of mixing of the air decreases the NO₂ and CO potential absorption.

References

- Hastie T.J., Tibshirani R.J.(1990) *Generalized Additive Models*. Chapman & Hall/CRC, New York.
- Holland D. M., De Oliveira V., Cox L. H., Smith R. L. (2000). Estimation of regional trends in sulfur dioxide over the eastern United States. *Environmetrics*; (11), 373-393
- Li, K., Le, N., SUN, L., and Zidek, J. V. (1999). Spatial-temporal models for ambient hourly PM10 in Vancouver. *Environmetrics*, (10), 321–338.