ISSN: 2037-7738
GRASPA Working Papers
[online]

Web Working Papers
by
The Italian Group of Environmental Statistics

Gruppo di Ricerca per le Applicazioni della Statistica
ai Problemi Ambientali

www.graspa.org

Air quality impact assessment of traffic policy in Milan

Alessandro Fassò

GRASPA Working paper n.46, September 2012

Air quality impact assessment of traffic policy in Milan

Alessandro Fassò
DIIMM - University of Bergamo

September 14, 2012

Abstract

This paper presents a general spatio-temporal model for assessing the air quality impact of environmental policies which are introduced as abrupt changes. The estimation method is based on the EM algorithm and the model allows to estimate the impact on air quality over a region and the reduction of human exposure following the considered environmental policy. Moreover, impact testing is proposed as a likelihood ratio test and the number of observations after intervention is computed in order to achieve a certain power for a minimal reduction. An extensive case study related to the introduction of the congestion charge in Milan city and the monitoring of particulate matters and total nitrogen oxides motivates the methods introduced and illustrate implementation issues and inferential machinery.

1 Introduction

Environmental, energetic and industrial policies are often motivated by the need to improve air quality in terms of pollutants concentration reduction. This is usually pursued by a supposedly appropriate cut of emissive activity like selective or indistinct car traffic reduction, carbon energy substitution with green energy, industrial requalification, etc. Hence a fundamental step is to assess if a policy actually obtains a relevant reduction of pollutant concentrations. In this paper, we develop a general observational methodology for spatiotemporal impact assessment.

The general approach proposed is motivated by and benchmarked on a real application related to the impact of traffic reduction. In particular, on January 16, 2012, the Municipality of Milan introduced a new traffic restriction system known as congestion charge. This requires to pay a fee of five Euros for entering the central area, known as "*Area C*" which is inside the so called "*Cerchia dei Bastioni*". According to Municipality, in the first two months, car traffic decreased of about 36% in Area C and, at the overall city level, Municipality reported a traffic reduction of about 6% which started at the beginning of January, before the congestion charge. This bringing forward may partly be related to the preparatory campaign playd out by the administration and partly by overall decrease in gas consumption at the national level due to economic crisis.

January and February have been cold and heavily polluted months with a large number of days exceeding the thresholds fixed by the European regulations (see Arduino and Fassò, 2012). Although the congestion charge is intended as a traffic control measure, the question which arise is whether there has been an impact on air quality or not. Air quality is usually defined according to the concentrations of various pollutants. In particular particulate matters (PM) are important because of their toxicity and are extensively monitored. Nevertheless the health effects of PM are known to depend not only on concentration but also on particle size, composition and black carbon content, so that having extensive data on particle numbers (see e.g. Hong-di and Wei-Zehn, 2012) or on black carbon content (see e.g. Janssen et al. 2011) would be very useful to understand air quality from health protection point of view. For other pollutants threatening human health, e.g. nitrogen oxides, which also are extensively monitored because of their health impact, we do not have this problem and concentration is the main health concern.

In general, comparing concentrations before and after intervention is a non-trivial problem. In fact, considering PM_{10} concentration in Milan city center depicted in Figure 1, we note that the January-February mean of the last three years is 71.3 $\mu g/m^3$ which is smaller than the after-intervention mean, obtained between January 16 and the end of February, which is 76.7 $\mu g/m^3$. So the

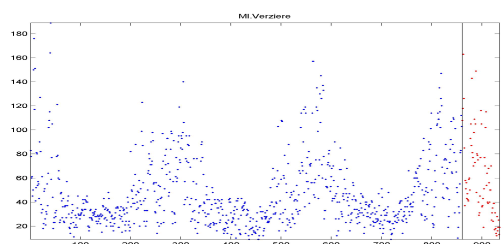

Figure 1: PM_{10} at Verziere Station. Black vertical line is Jan 16.

air quality impact of the traffic reduction due to the congestion charge has to face a major intrinsic confounder, due to atmospheric conditions, which greatly affects air quality in the Po Valley basin in general and Milan area in particular. Moreover a second possible confounder is related to the economic crisis which is reducing car use and gas consumption around Italy.

This paper aims at an early estimation of the spatially distributed reduction in the yearly average of pollutant concentrations. To do this, it aims to focus the following three scientific questions, which are exemplified by means of Milan case study.

1. The first point is whether or not the congestion charge has a measurable permanent impact on air quality in terms of concentration reduction of particulate matters (PM_{10} and $PM_{2.5}$) and total nitrogen oxides (NO_X) after adjusting for the effects of meteorological conditions and traffic reduction due to economic crisis or other general fluctuations which are not related to Milan specific facts.

2. Moreover, the second point aims to know if there is a different impact inside the intervention area (area C) and the rest of the city.

3. Finally, the third point is related to the spatial and temporal information content required to have sound conclusions. That is the number of days, required to "observe" a statistically significant permanent impact with high probability and the number of stations required to understand the spatial impact.

In time series analysis the non spatial part of questions one and two are treated by intervention analysis after the celebrated paper of Box and Tiao (1975) . See also e.g. Hipel and McLeod (2005). Soni et al. (2004) discuss spatiotemporal intervention analysis in the context of neurological signal analysis using STARMA models. In river networks water quality monitoring, Clement et al. (2006) considered a spatiotemporal model based on directed acyclic graphs. Here, we extend these methods to a general multivariate spatiotemporal model for air quality and develop some examples related to Milan congestion charge.

With this aim, the rest of the paper is organized as follows. Section 2 presents a general spatiotemporal model, which is capable of various levels of complexity according to the information content of the underlying monitoring network. The estimation method is based on the EM algorithm and the model allows to estimate the impact on air quality and the reduction of human exposure following the considered environmental policy. Moreover, impact testing is proposed as a likelihood ratio test and the number of observations after intervention is computed in order to achieve a certain power for a minimal reduction. In section 3, the above approach is applied to the introduction of the congestion charge in Milan city. To do this, the general model is tailored to the reduced monitoring network that the environmental agency, ARPA Lombardia, implemented for monitoring particulate matters (PM_{10} and $PM_{2.5}$) and nitrogen oxides (NO_X) in this city. The concentration reduction is then assessed for the above pollutants using a preliminary vector autoregression approach and a confirmatory spatiotemporal model named STEM used when the spatial information contained in the monitoring network is sufficient. The conclusions and acknowledgment sections close the paper.

2 Impact modelling by STEM

We consider here a general model able to assess the impact on air quality of an environmental rule in a geographic region $\mathcal{R}$. To do this, we define a spatiotemporal model for the observed concentration at coordinates $s \in \mathcal{R}$ and day

$t = 1, 2, ..., n$, denoted by $y(s,t)$, able to capture the effect of the environmental intervention, dated $t^* = n - m + 1$ and observed for $m = n - t^* + 1$ days, namely

$$y(s,t) = -\alpha(s,t) + \beta(t)\, x(s,t) + \zeta(s,t) \tag{1}$$

The quantity $\alpha(s,t)$ represents the expected spatial impact of the environmental intervention and $\alpha(s,t) = 0$ for $t < t^*$. In general terms α is a dynamic random field and the effectiveness of a environmental policy can be assessed by the expected impact on pollution concentrations over region $\mathcal{R}$ and time horizon M, which is given by

$$\Delta = \frac{1}{M} \sum_{t=t^*}^{t^*+M-1} \int_{\mathcal{R}} E(\alpha(s,t))\, p(s,t)\, ds$$

The weighting function $p(s,t)$ may be used for averaging, e.g. $p(s,t) = |\mathcal{R}|^{-1}$, or risk assessment. For example, we may be interested in human exposure and, following Finazzi et al (2012), we may take the weighting function $p(s,t)$ as the dynamic population distribution or a time-invariant $p(s)$ as a static population distribution over the study area. If $\Delta < 0$ then the impact is negative and we have an increase in pollutant concentration.

The simplest model for reduction assessment, is given by a scalar deterministic impact

$$\alpha(s,t) = \alpha \tag{2}$$

which assumes constant impact over time and space after intervention. At an intermediate complexity level, we may use $\alpha(s,t) = \alpha(s)$ which gives a time-invariant reduction map, appropriate for assessing a localized permanent stationary impact. Of course the choice among the above alternatives relies also on the spatial information content of the monitoring network.

Confounders may be covered by a time varying linear confounder model component $\beta(t)\, x(t)$, where $\beta(t)$ is a stationary stochastic coefficient vector. For example, Finazzi and Fassò (2011b) use a Markovian dynamics for $\beta(t)$. In the Milan application of the next section, considering the limited amount of spatiotemporal information contained in the data, in order to avoid overfitting and shadowing of the change point t^*, we use a deterministic β with a minimal seasonal structure given by:

$$\beta(t) = \begin{cases} \beta_s & t \in Summer \\ \beta_w & t \in Winter \end{cases} \tag{3}$$

In equation (1), the spatiotemporal error $\zeta(s,t)$ allows for spatial and temporal correlation using either separability or non-separability, see e.g. Porcu et al (2006), Bruno et al. (2009) and Cameletti et al. (2011). In this paper, we use a latent process with three components adapted from Fassò and Finazzi (2011a), namely

$$\zeta(s,t) = z(t) + \omega(s,t) + \varepsilon(s,t)$$

where $z(t) = \gamma z(t-1) + \eta(t)$ is a stable Gaussian Markovian process, with $|\gamma| < 1$ and $\sigma_\eta^2 = Var(\eta)$. The purely spatial component $\omega(s,t)$ is given by iid time replicates of a zero mean Gaussian spatial random field characterized by a spatial covariance function given by

$$\gamma(|s-s'|) = \sigma_\omega^2 \exp\left(-\frac{|s-s'|^r}{\theta}\right) \tag{4}$$

with $r = 1$ or 2. Finally $\varepsilon(s,t)$ is a Gaussian measurement error iid over time and space, with variance σ_ε^2.

2.1 Estimation and inference

We denote the parameter array characterizing model (1) by $\theta = (\theta_\alpha, \theta_{\sim\alpha})$, where θ_α is the component related solely to the effect $\alpha(\cdot) = \alpha(\cdot|\theta_\alpha)$ and $\theta_{\sim\alpha}$ is the parameter component for the global dynamics of y independent on the intervention. Although, in general the impact could depend on $\theta : \Delta = \Delta(\theta)$, it is convenient to define models for which Δ depends solely on θ_α. In the simple case of equation (2), we have $\alpha = \theta_\alpha \propto \Delta(\alpha)$.

With this notation, the estimated model parameter array is given by $\hat{\theta} = \left(\hat{\theta}_\alpha, \hat{\theta}_{\sim\alpha}\right)$, which may be computed using maximum likelihood as in Fassò and Finazzi (2011a). In particular the estimates are computed using the EM algorithm, hence the acronym STEM for this approach. Note that the EM algorithm relies on a posteriori common latent effects, namely $\hat{z}(t) = E(z(t)|Y)$, which are computed by the Kalman smoother, and a posteriori local effects, namely $\hat{\omega}(s,t) = E(\omega(s,t)|Y)$, which are computed by Gaussian conditional expectations. An efficient software for EM estimation, filtering and kriging, called D-STEM, has been recently introduced by Finazzi (2012) and is largely used in section 3.

Using the above model, the effectiveness of an environmental measure may be proven by rejecting the non-change hypothesis given by

$$H_0 : \Delta(\theta_\alpha) = 0$$

Suppose that $\Delta(\theta_\alpha) = 0$ for $\theta_\alpha = 0$. We can then test the above non effect hypothesis by the (one sided) likelihood ratio test. In particular, if $\theta_\alpha = \alpha$ is a simple scalar parameter, we can approximate the likelihood ratio test by a simple t test and reject H_0 for large significant values of $\frac{\hat{\alpha}}{se(\hat{\alpha})}$. Moreover, if $\alpha(s)$ is a stochastic map, it may be estimated by kriging-like computations, giving the impact map $\hat{\alpha}(\cdot)$ and the $1-p$ level confidence bands given by

$$\hat{\alpha}(\cdot) \pm z_{p/2} se(\hat{\alpha}(\cdot))$$

2.2 Days for detection and Information

Under regularity assumptions, the observed Fisher information matrix I_n, for large n and m, may be related to the partitioned information i_θ as follows

$$I_{n,\theta} \cong m \begin{pmatrix} i_\alpha & i_{\alpha,\sim\alpha} \\ i_{\sim\alpha,\alpha} & \frac{n}{m} i_{\sim\alpha} \end{pmatrix} \tag{5}$$

where the information blocks are conformable to $\theta = (\theta_\alpha, \theta_{\smile\alpha})$. It follows that the precision in the estimation of the pollution reduction α depends mainly on m.

The number of days required to detect a reduction of size α^* with high probability π is then computed with formulas which generalize the classical sampling results. In particular for the simple scalar parametrization of $\alpha(\cdot)$, applying the partitioned matrix inversion lemma to expression (5), we have the approximated formula

$$m \geq m^* = \left(i_\alpha^{-1} - i_{\alpha,\sim\alpha} \frac{m}{n} i_{\sim\alpha}^{-1} i_{\sim\alpha,\alpha} \right)^{-1} \left(\frac{\Phi(1-p)^{-1} + \Phi(\pi)^{-1}}{\alpha^*} \right)^2 \tag{6}$$

3 Milan case study

The monitoring network of Milan city, depicted in Figure 2, is composed by one station for $PM_{2.5}$, four stations for PM_{10} and eight stations for NO_X. All these sensors give data at least three years old, that is, we consider the log-transformed and centered concentrations between January 1^{st}, 2009, and July 20, 2012, totalling $n = 1297$ days and $m = 186$ days are after the congestion charge introduction dated January 16, 2012.

In order to adjust for the confounders, we considered meteorological conditions in Milan, namely wind speed (daily maximum and average) and direction, humidity, temperature, solar radiation and pressure. Moreover, we used the additional covariates given by the concentration readings of the same pollutants in Bergamo city which is approximately 45km North-East of Milan and we used data from the urban ground station known as Meucci. This is an important proxy for all other meteorological and economic factors common to the north plain of Lombardy but the congestion charge in Milan.

3.1 Modelling details

Due to the sparsity of Milan data, we start using three different simple exploratory q-dimensional vector seasonal autoregressive (ARX) models. For $PM_{2.5}$ we use only a unidimensional ARX model, while for PM_{10} and NO_X, we integrate the analysis of the vector ARX models with the STEM approach of the previous section.

The exploratory ARX models are given by

$$y_t = -a_t + bx_t + Gy_{t-1} + e_t$$

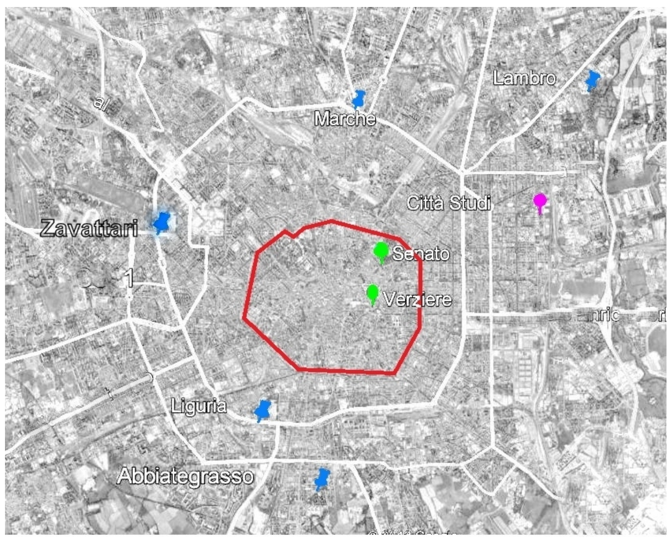

Figure 2: Milan Area C (red line) and monitoring network.
Blue pushpins: NO_X only, from top right counterclockwise Lambro, Marche, Zavattari, Liguria, Abbiategrasso.
Green Ballons: PM_{10} and NO_X, Area C, Senato and Verziere.
Pink Ballon: $PM_{2.5}$, PM_{10} and NO_X . Città studi, Pascal.

where y_t is $PM_{2.5}$, PM_{10} or NO_X with dimensions $q = 1, 3$ and 7, respectively. The impact a_t is as in equation (2) and, similarly, the seasonal adjustment b is as in equation (3). The covariates are selected by taking into account BIC criterion and residual autocorrelation. The persistence matrix G is taken as a diagonal matrix:

$$G = diag\left(g_1, ..., g_q\right)$$

Note that the multivariate approach is important here, because the errors e_t may be strongly (spatially) correlated.

According to the AR(1) dynamics, the scalar steady state impact on y_t is given by

$$d = \frac{a}{1-g} \tag{7}$$

Moreover, ignoring the uncertainty of the pre-intervention estimation of g and recalling the Fisher information matrix given in (5), we get the approximate variance for $\hat{d}$, namely

$$Var\left(\hat{d}\right) \cong Var\left(\hat{a}\right) / \left(1-\hat{g}\right)^2$$

It follows that the city average steady state effect is given by

$$\bar{d} = \Sigma_{j=1}^{q} \hat{d}_j p_j$$

with variance given by the well known quadratic form:

$$Var\left(\bar{d}\right) = p' Var\left(\hat{D}\right) p$$

In the above formula, $D = (d_1, ..., d_q)$ and the weights $p = (p_1, ..., p_q)$ can be based on the population density, which is taken as approximately constant in the rest of the paper, as we are involved in city center data especially for the PM study.

3.2 Fine particulate matters

We start with the single station on fine particulate matters $PM_{2.5}$, namely Pascal station, which is a ground station external to "Area C" and located in the relatively central quarter named "Città studi". Here, fine particulate concentrations have a three-year average of approximately $30\mu g/m^3$ before intervention. After January 16 the January-February three year mean increased from 53.7 $\mu g/m^3$ to 54.1 $\mu g/m^3$, questioning the congestion charge effect.

To fit the model for fine particulate matters, we use the centered log transformed concentrations which have variance $\sigma_y^2 = 0.70$, and, after row deletion of missing data, we get the fitted model of Table 1 based on the remaining 1038 observations. It is worth noting that wind speed has an effect both in summer and winter while, not surprisingly, solar radiation is not significant in winter. After adjusting for Bergamo concentrations which are quite significant in the model, the residual autocorrelation is not large resulting in $g = 0.212$.

Parameter	Estimate	se
a	0.095	0.031
$PM_{2.5}$ in Bergamo	0.775	0.027
Average Wind speed - Summer	-0.295	0.049
Max Wind speed - Summer	0.044	0.025
Solar radiation - Summer	-0.00065	0.0002
Average Wind speed - Winter	-0.312	0.063
Max Wind speed - Winter	0.062	0.033
g	0.212	0.021
d	0.121	0.040
m^*	162	
R^2	0.80	

Table 1: ARX model and reduction for $PM_{2.5}$

Since we are in log scale, the permanent effect, computed as both $\hat{a}$ or $\hat{d}$ of equation (7), may be interpreted as a percent change. According to this model, at Pascal station, we observe a 0.12 permanent reduction of $PM_{2.5}$ on log scale with a one-sided p-value smaller than 0.2%. From the bottom row of Table 1, we see that fitting is quite good. The residuals result to be satisfactorily white noise but moderately non Gaussian as shown by Figure 3 and by unreported kurtosis which are larger than three. Although some alternatives to conditional Gaussian models could be developed, see e.g. Bartoletti and Loperfido (2010) or Nadarajah (2008), the above results where validated by simulation experiments and by comparison with robust estimation methods getting very close results to Table 1.

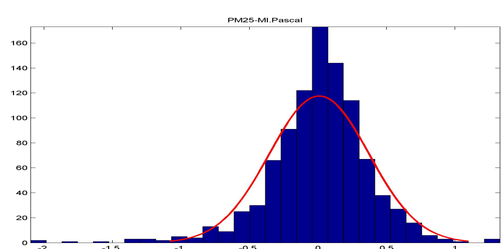

Figure 3: Residuals of $PM_{2.5}$ at Pascal station.

Days for detection From the practical point of view it is important to see which is the number of days required to detect a certain reduction in the annual average of $PM_{2.5}$. Following the approach which gives formula (6), we

Station	a	PM_{10} in Bergamo	Max WS Winter	Average WS Summer	Average WS Winter	g	d	R^2
Pascal	0.045	0.718	0.169	-0.360	-0.576	0.193	0.068	0.80
se	0.024	0.023	0.065	0.049	0.094	0.020	0.030	
Senato	0.059	0.675	0.150	-0.290	-0.428	0.210	0.075	0.80
se	0.022	0.021	0.061	0.046	0.088	0.020	0.030	
Verziere	0.069	0.625	0.237	-0.182	-0.504	0.224	0.089	0.85
se	0.018	0.017	0.049	0.037	0.070	0.018	0.023	
Milan	0.061						0.077	0.82
	0.018						0.024	

Table 2: Vector ARX model and reduction for PM_{10}, Single stations and city average. WS stands for Wind speed.

consider the t-test for the hypothesis of no impact $H_0 : d = 0$ against a reduction $d > 0$ based on large values of the statistic $\frac{\hat{a}}{se(\hat{a})}$. Using the nominal significance level $p = 5\%$, the number of days for detecting a permanent reduction of size $d^* = 0.10$ in log scale with probability $\pi = 85\%$ is given by $m \geq m^* = 162$ as shown in Table 1, which is consistent with the above results.

3.3 Particulate matters

For the PM_{10}, we have three sites, two are traffic stations located inside "Area C", namely Verziere and Senato stations, and the third is Pascal, which is a ground station located in Città studi which is semi-peripheral area with patterns similar to the city center.

Table 2 shows the fitted model for PM_{10}. We see that the persistence coefficients g are small and very close each other, denoting the same weak autocorrelation after adjusting for the Po valley concentration proxy given by Bergamo measurement and local meteorological covariates. Both average and maximum wind speed have an effect on PM_{10} with a clear seasonal behaviour. The last column shows that fitting is quite satisfactory. We note that, after introducing the Bergamo proxy, only local conditions given by wind speed enter as additional covariates.

As in the previous section, a moderate residual non normality indicated by kurtosis larger than three for all components does not jeopardize the results of Table 2. Moreover, Figure **??** shows that the residuals can be assumed to be a white noise and the importance of the multivariate approach is appreciated by the marked residual correlation shown in Table 3.

Note that the global three-year average for these stations before intervention is about $45\mu g/m^3$, so the average reduction of $d = 0.137$ in log scale, with $se = 0.053$, corresponds approximately to $6\mu g/m^3$ for the yearly average. According to the data, this impact is stronger in Area C.

Città studi	1	0.65	0.58
Senato		1	0.69
Verziere			1

Table 3: ARX model residual correlations for PM_{10}

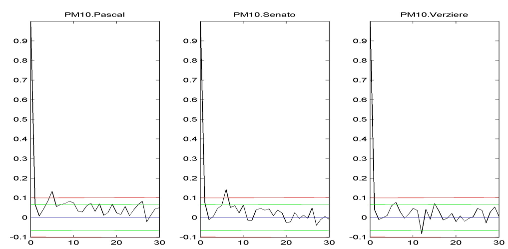

Figure 4: Residual autocorrelations of Vector ARX model for PM_{10}.

STEM approach In the light of the preliminary analysis, we go further with the approach of section 2, obtaining the fitted model of Table 4. To avoid initial values dependence, the EM algorithm has been replicated 100 times applying beta distributed random perturbations[1] to the initial estimates which have been computed using the method of moments. The power r of the spatial correlation (4) has been selected to $r = 1$ comparing the corresponding log-likelihoods.

Generally speaking the standard errors are small, with the exception of the spatial correlation parameter θ, which has a quite large uncertainty. This is not surprising since with only three stations is not easy to estimate spatial correlation. The reduction parameter α is positive, denoting a significant permanent reduction of particulate concentrations of 0.085 in log scale for the city center with one sided p-value smaller than 1%.

Moreover, the likelihood ratio test for the hypothesis of no effect, namely $H_0 : \alpha = 0$, gives a test statistic $2log\,(LR) = 13.2$ with a p-value smaller than 0.1%. Finally, the number of days for detection of formula (6), with nominal significance level $p = 5\%$, a permanent reduction of size $\alpha^* = 10\%$ and power $\pi = 85\%$ results to be $m^* = 154$ which is consistent with above results.

According to ARX model, Area C has a little larger PM_{10} reduction, but the difference in a coefficients of Table 2 is far from significant. Using STEM model, the corresponding analysis is based on the a posteriori local effects of section 2.1, that is $\bar{\omega}\,(s) = \frac{1}{m}\sum_{t=t^*}^{n} \hat{\omega}\,(s,t)$ which are smaller than 1% and have non significant p-values for Pascal, Senato and Verziere. Hence we can conclude the reduction in PM_{10} is approximately constant around the city center but we

[1] At each replication, the initial values have been rescaled between 0 and three times their value by multiplication with random numbers $3B$ where $B's$ have been drawn from the Beta distribution with parameters 4 and 8, so that $E\,(3B) = 1$.

Parameter	Estimate	se
α - City center impact	0.085	0.035
PM_{10} in Bergamo	0.790	0.014
Summer - Average Wind speed	-0.242	0.033
Winter - Average Wind speed	-0.536	0.062
Winter - Max Wind speed	0.286	0.046
γ	0.378	0.10
σ_η^2	0.032	0.011
θ	12.4	10.1
σ_θ	0.131	0.038
σ_ε^2	0.028	0.002
Log-likelihood	3073.6	

Table 4: STEM model for PM_{10}

have no information on peripheral areas.

3.4 Nitrogen oxides

As mentioned above, the total nitrogen oxides are monitored in Milan city more extensively than particulate matters. In fact, in addition to the previous three stations, we have four traffic stations near the internal bypass and one ground station in the green area of Parco Lambro, near the eastern circular highway and exposed to Linate airport emissions. Using this relatively larger network, we will address the spatial variability of the congestion charge impact around the city. We first observe that nitrogen oxides data are given as hourly data. In order to have high quality daily data, and considering that STEM approach is resistant to missing data, we defined as missing those daily averages based on less then 20 validated observations. An additional meteorological covariate enters at this stage which was not significant with particulate matters. This is the prevalence of South western wind (SW-PWD) defined as the number of hours per day when the prevailing wind is from SW.

The vector ARX model for these eight stations is reported in Table 5. We keep this as a preliminary analysis model even if, with such a number of spatial locations, a vector approach begins showing its limitations and the need for a unified approach such as STEM is now becoming evident. The proxy from Bergamo and the seasonal effect on wind are again decisive for the good model fitting, reported in the last column of Table 5. As mentioned above, the wind direction from South West is significant for various stations. Despite the number of monitoring stations, the impact of the congestion charge on NO_X is far from constant around the city.

In Pascal station, we have the maximum reduction, more than 0.28 in log scale and p-value close to zero. Surprisingly this value is much larger than inside the traffic restricted Area C, where the permanent effect estimated by the ARX model is positive but very small especially in Verziere. Moreover, in peripheral stations, at Liguria and Lambro stations, we have significant increases in NO_X concentrations. Note that this fact is not a model artifact but reflects the daily averages behaviour. These points will be partially overcome in the final STEM model.

Station	a	Summer			Winter		NO_X in Bergamo	g	R^2
		Ave WS	Max WS	SW-PWD	Ave WS	Max WS			
Pascal	0.295	-1.004	0.401	-0.0017	-1.20	0.487	0.539	-0.328	0.85
se	0.03	0.12	0.09	0.003	0.13	0.10	0.03	0.02	
Senato	0.046	-0.303	0.107	-0.0077	-0.618	0.335	0.396	-0.420	0.87
se	0.025	0.094	0.072	0.0026	0.11	0.084	0.024	0.025	
Verziere	-0.0064	-0.669	0.323	-0.0095	-0.839	0.376	0.489	-0.260	0.83
se	0.025	0.097	0.074	0.003	0.11	0.087	0.025	0.025	
Liguria	-0.136	-0.760	0.390	-0.022	-0.779	0.345	0.307	-0.360	0.79
se	0.02	0.091	0.069	0.003	0.10	0.081	0.02	0.02	
Marche	0.018	-0.75	0.228	-0.023	-0.96	0.33	0.340	-0.256	0.83
se	0.02	0.09	0.06	0.002	0.10	0.08	0.02	0.02	
Abbiategrasso	-0.054	-0.807	0.288	-0.0011	-1.18	0.630	0.400	-0.368	0.85
se	0.03	0.11	0.08	0.003	0.12	0.097	0.027	0.025	
Lambro	-0.115	-0.736	0.288	-0.0054	-1.17	0.552	0.374	-0.365	0.81
se	0.03	0.10	0.08	0.0029	0.12	0.093	0.026	0.025	
Zavattari	-0.013	-0.701	0.214	0.017	-0.98	0.470	0.454	-0.239	0.83
se	0.03	0.087	0.07	0.002	0.099	0.078	0.023	0.024	

Table 5: Vector ARX model for NO_X. WS stands for Wind Speed, SW-PWD for South western prevailing wind direction.

The residuals of this model are satisfactorily Gaussian, white noise and homoskedastic according to standard tests. Note that the estimated permanent reduction for Pascal is quite large, $\hat{d} = 0.44$ with a standard error $se = 0.05$, and will be discussed in the next section.

STEM approach In order to deepen the urban variability and seek for a unified conclusion about the congestion charge impact on nitrogen oxides concentrations, we consider a STEM model with spatially varying impact. In the light of the limited spatial information contained in Milan eight station network, we use the following simple impact model for $t \geq t^*$

$$\alpha(s) = \begin{cases} \alpha_1 & s \in \text{City center} \\ \alpha_2 & s \notin \text{City center} \end{cases} \tag{8}$$

where, for the purpose of this paper, the city center is defined by Area C plus Città studi, as discussed in the previous section. The estimated model, reported in Table 6, clearly shows the difference between the impact in the city center, where a permanent reduction of about 0.23 in log scale is estimated, and the peripheral area where NO_X concentrations show an increase of 0.07, although not statistically significant. To have a confirmation about the non-increase in the peripheral area, we tested the hypothesis $\alpha_2 = 0$ using the Likelihood ratio test. This approach gives a restricted Log-likelihood of 6515.0 and a non significant LR statistic with p-value = 15.7%. Interestingly the spatial correlation parameter θ is very close to the PM_{10} case but, as expected, the standard deviation is smaller reflecting the major spatial information of nitrogen oxides network.

The local effects $\bar{\omega}(s)$ of section 2.1 are reported in Table 7. Note that, considering for example Pascal, if we sum up α_1 from Table 6 and $-\bar{\omega}(s)$ from Table 7 we get a total reduction of 0.402 which is quite close to the estimate $\hat{d}$ of the ARX model. Note that these results reflect the change in the unadjusted seasonal average concentration. For example in the last two columns of Table 7, we have the comparison of the average concentrations in the period January 16 - July 20 before and after the congestion charge introduction, in the years 2009-2011 and 2012 respectively.

4 Conclusions

In order to answer the three scientific questions on air quality impact raised in the introduction, we introduced a general approach, based on STEM model, for spatiotemporal impact assessment of air quality policies allowing both estimation and testing.

The first question on the presence of a "permanent impact" is positive. In particular, we showed that the congestion charge operating in Milan center since January 16, 2012, has a significant permanent impact on air quality in terms

Parameter		Estimate	se
Central impact - α_1		0.218	0.048
Peripheral impact - α_2		-0.077	0.047
Summer	Average WS	-0.706	0.054
	Max WS	0.267	0.042
	South West	-0.0073	0.0014
Winter	Average WS	-1.021	0.061
	Max WS	0.552	0.050
BG.NO_X		0.604	0.016
γ		0.599	0.048
σ_η^2		0.033	0.0047
θ		12.7	3.7
σ_θ		0.153	0.013
σ_ε^2		0.054	0.002
Log-likelihood		6516.0	

Table 6: STEM model for NO_X. WS stands for Wind speed.

Station	Local Effect	se	Seasonal average	
			before	after
Pascal	-0.160	0.065	56.3	31.3
Verziere	0.138	0.064	67.1	51.4
Senato	0.117	0.064	46.5	41.2
Liguria	0.157	0.065	63.2	73.4
Marche	-0.070	0.065	70.2	67.2
Abbiategrasso	0.100	0.065	33.4	42.4
Lambro	-0.065	0.066	51.4	56.8
Zavattari	0.035	0.066	65.5	61.3

Table 7: STEM local effects $\bar{\omega}(s)$ for NO_X, in log scale, and unadjusted seasonal averages before(2009-2011) and after (2012) the intervention in *ppb*.

of particulate matters and nitrogen oxides concentrations at least in the city center. The air quality impact has been estimated after adjusting for meteorological factors and other common forcing factors, such as the economic crisis. Interestingly, the reduction on $PM_{2.5}$, PM_{10} and NO_X concentrations estimated using both a preliminary vector autoregressive model and STEM approach, is not confined inside the traffic restricted area.

The second question, on the spatial distribution of the air quality change has an articulated answer. We observed that, despite the reduced number of monitoring stations in the city, the impact has a noticeable spatial variability which is different for PM_{10} and NO_X. In particular, after the traffic intervention, a significant reduction of both particulate matters and nitrogen oxides has been estimated in the city center. The reduction of particulate matters, which is about 8%, or 3.6 $\mu g/m^3$, in city center, is slightly higher in the intervention area, but the spatial variations around the city center either inside or outside the Area C can be neglected. The nitrogen oxides show different figures and pattern, with a reduction larger than 19%, or 13 *ppb*, in city center. Surprisingly, the reduction is higher in Città studi, outside the intervention area, while in Verziere, which is inside Area C, the reduction is barely significant. Moreover, in the peripheral areas of the city, we observe changes of nitrogen oxides in both directions and no overall decrease can be concluded.

Considering the third question on spatiotemporal information, we observe that, using the data before August 2012 gives a substantially clear picture of air quality impact, at least in the city center. Nevertheless, having a more extended monitoring network, would allow us to estimate more detailed reduction maps both for pollutant concentrations and human exposure by means of the dynamical kriging capabilities of the STEM approach proposed in this paper.

5 Acknowledgments

This research is part of Project EN17, 'Methods for the integration of different renewable energy sources and impact monitoring with satellite data', Lombardy Region under 'Frame Agreement 2009'.

The help of Angela Locatelli and Francesco Miazzo for data handling has been much appreciated by the author.

References

[1] Arduino G., Fassò A., (2012) Environmental regulation in the European Union. In Abdel H. El-Shaarawi and Walter W. Piegorsch (eds) Encyclopedia of Environmetrics 2nd ed., Wiley. In printing.

[2] Bartoletti S., Loperfido N. (2010) Modelling air pollution data by the skew-normal distribution. *Stochastic Environmental Research and Risk Assessment.* 24:513–517.

[3] Box G. E. P. and Tiao G. C. (1975) Intervention Analysis with Applications to Economic and Environmental Problems. Journal of American Statistical Association. 70(349):70-79.

[4] Bruno F., Guttorp ·P., Sampson P.D., Cocchi D. (2009) A simple non-separable, non-stationary spatiotemporal model for ozone. *Environ Ecol Stat.*16:515–529.

[5] Cameletti M. Ignaccolo R. Bande S. (2011) Comparing spatio-temporal models for particulate matter in Piemonte, *Environmetrics.* 22(8):985–996

[6] Clement L., Thas O., Vanrolleghem P.A. and Ottoy J.P. (2006) . Spatio-temporal statistical models for river monitoring networks. Water Science & Technology. 53(1):9–15.

[7] Fassò A, Finazzi F, (2011a) Maximum likelihood estimation of the dynamic coregionalization model with heterotopic data. *Environmetrics.* 22(6):735-748.

[8] Finazzi F. (2012) D-STEM – A statistical software for multivariate space-time environmental data modeling. METMA VI International Workshop on spatio-temporal modelling. Guimaraes, 12-14 September 2012. In printing.

[9] F. Finazzi, A. Fassò (2011b). Spatio temporal models with varying coefficients. In Cafarelli Ed. (2011) Spatial Data Methods for Environmental and Ecological Processes – 2nd Edition, Foggia, Sept. 1-2, 2011. ISBN: 978-88-96025-12-3. Also in: *Graspa WP*, 2011, www.graspa.org, ISSN: 2037-7738.

[10] Finazzi F., Scott M.E. , Fassò A. (2012). A statistical framework for model based air quality indicators and population risk evaluation. *GRASPA WP*, 43:1-26. www.graspa.org, ISSN: 2037-7738.

[11] Graf-Jaccottet M, Jaunin MH. (1998) Predictive models for ground ozone and nitrogen dioxide time series. *Environmetrics.* 9:393–406.

[12] Hipel, K.W. & McLeod, A.I. (2005) *Time Series Modelling of Water Resources and Environmental Systems.* © 2005 Hipel, K.W. & McLeod, A.I.. http://www.stats.uwo.ca/faculty/aim/1994Book/

[13] Hong-di H., Wei-Zhen L. (2012) Urban aerosol particulates on Hong Kong roadsides: size distribution and concentration levels with time. *Stochastic Environmental Research and Risk Assessment.* 26:177–187.

[14] Janssen N.A., Hoek G., Simic-Lawson M., Fischer P., van Bree L., ten Brink H., Keuken M., Atkinson R.W., Anderson H.R., Brunekreef B., Cassee F.R. (2011) Black Carbon as an Additional Indicator of the Adverse Health Effects of Airborne Particles Compared with PM10 and PM2.5. *Environmental Health Perspectives.* 119:1691-1699.

[15] Nadarajah S (2008) A truncated inverted beta distribution with application to air pollution data. *Stochastic Environmental Research and Risk Assessment.* 22:285–289.

[16] Porcu E., Gregori P., J Mateu J. (2006) Nonseparable stationary anisotropic space–time covariance functions. *Stochastic Environmental Research and Risk Assessment.* 21:113-122

[17] Soni P, Chan Y, Preissl H, Eswaran H, Wilson J, Murphy P, Lowery CL (2004) Spatial-Temporal Analysis of Non-stationary fMEG Data. *Neurology and Clinical Neurophysiology.* 100:1-6.