

# Statistical assessment of air quality interventions

Alessandro Fassò

Published online: 21 March 2013  
© Springer-Verlag Berlin Heidelberg 2013

**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 is related to the introduction of the congestion charge in Milan city. The consequent estimated reduction of airborne particulate matters and total nitrogen oxides motivates the methods introduced while its derivation illustrates both implementation and inferential issues.

**Keywords** Spatiotemporal models · Spatiotemporal intervention analysis · Air quality monitoring · Congestion charge · Particulate matters · Nitrogen oxides · STEM

## 1 Introduction

Environmental, energy 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, for example reduction of car traffic, substitution of carbon with green energy or requalification of heavy industry. Hence a fundamental step is to assess if a given policy actually

obtains a relevant reduction of pollutant concentrations. In this paper, we develop a general statistical 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 drivers 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 by 36 % in Area C and, at the overall city level, Municipality reported a traffic reduction by 6 % which started at the beginning of January, before the congestion charge. This preemptive reduction may be partly due to the preparatory campaign played out by the administration and partly to an overall decrease in gas consumption at the national level caused by the economic crisis.

January and February were 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 measure of traffic control, the question which arises is whether there has been an impact on air quality or not. Air quality is usually defined according to the concentrations of various pollutants entering a suitable air quality index, see e.g. Bruno and Cocchi (2007). We will focus here on two important compounds entering most air quality indexes, namely particulate matters and nitrogen oxides which are important for their toxicity. Particulate matters are predominantly secondary pollutants. Hence they have a background level which is large in percentage and difficult to reduce through local traffic policies. Moreover their health effects are known to depend not only on their concentration

---

A. Fassò (✉)  
DIIMM, University of Bergamo, Bergamo, Italy  
e-mail: alessandro.fasso@unibg.it

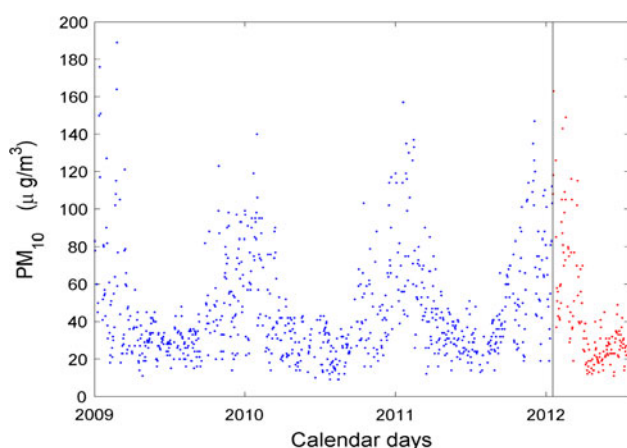
but also on particle size, composition and black carbon content, therefore, having extensive data on particle numbers (see e.g. Hong-di and Wei-Zehn 2012) or on black carbon content (see e.g. Janssen et al. 2011) would be very useful to understand air quality from a health protection point of view. To a large extent, total nitrogen oxides are compounds of primary pollutants. Hence they are closely linked to local traffic and have a background level which is smaller in percentage than particulate matters.

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 Fig. 1, we note that the January–February mean of the last three years is  $71.3 \mu\text{g}/\text{m}^3$  which is smaller than the after-intervention mean, obtained between January 16 and the end of February, which is  $76.7 \mu\text{g}/\text{m}^3$ .

Hence the air quality impact of the congestion charge may be obscured by atmospheric conditions, which are known to greatly affect the air quality of the Po Valley basin in general and the Milan area in particular. Moreover a second possible confounder is related to the economic crisis which is reducing both 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 on the following three scientific questions, which are exemplified by means of the Milan case study.

1. The first point is whether or not the congestion charge has a measurable permanent impact on air quality after adjusting for meteorological conditions and traffic reduction due to economic crisis.



**Fig. 1**  $PM_{10}$  at Verziere Station. Black vertical line is the intervention date on Jan 16, 2012

2. Moreover, the second point aim is to know if there is a different impact within 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 spatio-temporal 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 air quality model and develop some examples related to Milan congestion charge. In principle intervention analysis for air quality time series could also be based on neural networks, for example extending Corani (2005) or Gardner and Dorling (1999). Nevertheless this approach does not exploit the spatial dimension in the same explicit way and it is not further considered here.

The rest of the paper is organized as follows. We begin Sect. 2 with a preliminary vector autoregressive model and develop the section with a general spatiotemporal model, which is capable of various levels of complexity according to the information content of the underlying monitoring network. The model allows us to estimate the impact on air quality and the reduction of human exposure following the considered environmental policy. Maximum likelihood estimation is obtained by a version of the EM algorithm and impact testing is proposed as a likelihood ratio test. Moreover we give formulas for the minimum number of observations after intervention guaranteeing a certain high probability of detecting a prespecified hypothetical reduction of pollution. In Sect. 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 the city. The concentration reduction is then assessed for the above pollutants using a preliminary approach based on vector autoregressions and a final spatiotemporal model named STEM used when the spatial information contained in the monitoring network is sufficient. Section 4 discusses the results and gives concluding remarks to the paper, which is closed by an acknowledgment section.

## 2 Impact modelling

In this section we discuss the extension of intervention analysis to spatiotemporal data. To begin with, we introduce impact modelling in the simple case where data are collected by a single station, and no spatial correlation arises, or by a limited number of stations, and spatiotemporal correlation is modelled by a vector autoregression, see e.g. Hipel and McLeod (2005). This approach is suggested here only as a preliminary analysis to be integrated by a full spatio-temporal model as discussed in Sect. 2.2

### 2.1 Autoregressive approach

Let  $y(s_j, t)$  be the observation of a pollutant concentration from station located in  $s_j, j = 1, \dots, q$  and time  $t = 1, \dots, n$ . Moreover, let  $y_t = (y(s_1, t), \dots, y(s_q, t))$  be the vector of all concentrations at time  $t$  and, similarly, let  $x_t$  be the covariate vector including trends and confounders. Hence the exploratory intervention model is given by

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

In this model  $e_t$  is a Gaussian white noise and the impact  $a_t$  is a step function vanishing before intervention time  $t^*$ . For  $t \geq t^*$  we have  $a_t = a$  which is a  $q$  – dimensional vector whose positive elements define the pollution reduction size at the corresponding location  $s_j, j = 1, \dots, q$ . The persistence matrix  $G$  is taken as a diagonal matrix:

$$G = \text{diag}(g_1, \dots, g_q).$$

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

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

Moreover, ignoring the uncertainty of the pre-intervention estimation of  $g$ , we get the approximate variance for  $\hat{d}$ , namely

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

It follows that the city average steady state effect is given by  $\bar{d} = \sum_{j=1}^q \hat{d}_j p_j$  with variance given by the well known quadratic form  $\text{Var}(\bar{d}) = p' \text{Var}(\hat{D}) p$ . In the last formula,  $\hat{D} = (\hat{d}_1, \dots, \hat{d}_q)$  and the weights  $p = (p_1, \dots, p_q)$  can be based on the population density.

### 2.2 STEM approach

Following the modern hierarchical modeling approach to spatiotemporal data, we consider here a general model able

to assess the impact on air quality of an environmental rule in a geographic region  $\mathcal{D}$ . To do this, we define a spatio-temporal model for the observed concentration at coordinates  $s \in \mathcal{D}$  and day  $t = 1, 2, \dots, n$ , 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

$$y(s, t) = -\alpha(s, t) + \beta(t)x(s, t) + \zeta(s, t). \tag{2}$$

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

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

The weighting function  $p(s, t)$  may be used for averaging, e.g.  $p(s, t) = |\mathcal{D}|^{-1}$ , or risk assessment. For example, we may be interested in human exposure and, following Finazzi et al. (2013), 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

$$\alpha(s, t) = \alpha \tag{3}$$

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 linear confounder model component  $\beta x(t)$  or, more generally, by a time varying linear component  $\beta(t) x(t)$ , where  $\beta(t)$  is a deterministic varying coefficient vector or a stochastic process. For example, Finazzi and Fassò (2011) use Markovian dynamics for  $\beta(t)$ .

In Eq. (2), 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, in the light of limited spatial information available, we use a simple additively separable latent process with three components adapted from Fassò and Finazzi (2011), namely

$$\zeta(s, t) = z(t) + \omega(s, t) + \varepsilon(s, t) \quad (4)$$

where  $z(t) = \gamma z(t-1) + \eta(t)$  is a stable Gaussian Markovian process, with  $|\gamma| < 1$  and  $\sigma_\eta^2 = \text{Var}(\eta)$ . The purely spatial component  $\omega(s, t)$  is given by iid time replicates of a zero mean Gaussian spatial random field characterized by a suitable spatial covariance function  $\gamma(|s - s'|)$ . Finally  $\varepsilon(s, t)$  is a Gaussian measurement error iid over time and space, with variance  $\sigma_\varepsilon^2$ .

### 2.3 Estimation and inference

We denote the parameter array characterizing model (2) by  $\theta = (\theta_\alpha, \theta_{\sim\alpha})$ , where  $\theta_\alpha$  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  $\Delta$  depends solely on  $\theta_\alpha$ . In the simple case of equation (3), we have  $\alpha = \theta_\alpha \propto \Delta(\alpha)$ .

With this notation, the estimated model parameter array is given by  $\hat{\theta} = (\hat{\theta}_\alpha, \hat{\theta}_{\sim\alpha})$ , which may be computed using maximum likelihood as in Fassò and Finazzi (2011). 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 Sect. 3.

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

$$H_0: \Delta(\theta_\alpha) = 0. \quad (5)$$

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  $\frac{\hat{\alpha}}{se(\hat{\alpha})}$  where  $se(\hat{\alpha})$  is a suitable estimate of  $\sqrt{\text{Var}(\hat{\alpha})}$ . Moreover, if  $\alpha(s)$  is a Gaussian field, it may be estimated by kriging-like computations, giving the impact map  $\hat{\alpha}(\cdot)$  and the  $1 - p$  level approximate confidence bands given by

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

where  $z_p$  is the  $(1 - p)$  quantile of the standard Gaussian distribution.

### 2.4 Days for detection and information

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

$$I_{n,\theta} \cong m \begin{pmatrix} i_\alpha & i_{\sim\alpha,\alpha} \\ i_{\alpha,\sim\alpha} & \frac{m}{m} i_{\sim\alpha} \end{pmatrix} \quad (6)$$

where the matrix blocks  $i_\alpha, i_{\sim\alpha}$  and  $i_{\alpha,\sim\alpha}$  are the corresponding blocks of the non singular information  $i_\theta = -E\left(\frac{\partial^2}{\partial\theta\partial\theta'} \log(f(\theta))\right)$ . It follows that the precision in the estimation of the pollution reduction  $\alpha$  depends mainly on  $m$ .

Hence in testing (5) for  $\theta_\alpha = \alpha$ , the number of days required to detect a reduction of size  $\alpha^*$  with high probability  $\pi$  is then computed with formulas which generalise the classical sampling results. In particular for the simple scalar parametrization of  $\alpha(\cdot)$ , applying the well known matrix inversion lemma (see e.g. Gentle, 2007, 3.4.1) to expression (6), we have the approximated formula

$$m \geq m^* = \left( i_\alpha - \frac{m}{n} i_{\alpha,\sim\alpha} i_{\sim\alpha}^{-1} i_{\sim\alpha,\alpha} \right)^{-1} \left( \frac{z_p + z_{1-\pi}}{\alpha^*} \right)^2. \quad (7)$$

## 3 Milan case study

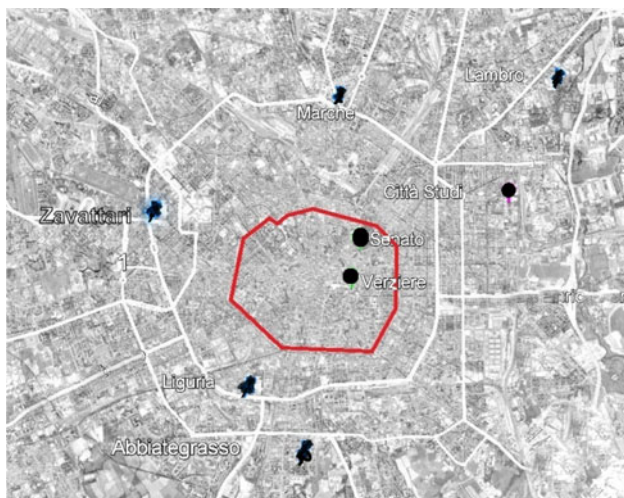
The monitoring network of Milan city, depicted in Fig. 2, is composed by one station for PM<sub>2.5</sub>, four stations for PM<sub>10</sub> and eight stations for NO<sub>x</sub>. We consider the log-transformed and centered concentrations between January 1st, 2009, and July 20, 2012, totalling  $n = 1297$  of which and  $m = 186$  after the introduction of the congestion charge, 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 from Meucci station in the city of Bergamo, which is approximately 45 km North-East of Milan. This is an important proxy for all other meteorological and economic factors which are common to the northern plain of Lombardy.

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  $\beta(t)$  with a minimal seasonal structure given by:

$$\beta(t) = \begin{cases} \beta_s & t \in \text{Summer} \\ \beta_w & t \in \text{Winter} \end{cases}. \quad (8)$$

Similarly the spatial covariance function of the spatial component  $\omega(s, t)$  in Eq. (4) is given by



**Fig. 2** Milan Area C (red line) and monitoring network. Pushpins: NO<sub>x</sub> only, from top right counterclockwise Lambro, Marche, Zavattari, Liguria and Abbiategrasso stations. Ballons: PM<sub>10</sub> and NO<sub>x</sub>, Area C, Senato and Verziere, Città studi, Pascal (including also PM<sub>2.5</sub>). (Color figure online)

$$\gamma(|s - s'|) = \sigma_\omega^2 \exp\left(-\frac{|s - s'|^r}{\rho}\right). \quad (9)$$

Although  $r$  could be estimated as a continuous parameter governing mean square differentiability of  $\omega$ , due to the limited spatial information, we calibrated this parameter only for  $r = 1$  or 2 and concluded that  $r = 1$  is better in terms of attained log-likelihood, both for particulate matters and nitrogen oxides.

### 3.1 Fine particulate matters

We start with the single station on fine particulate matters PM<sub>2.5</sub>, 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 μg/m<sup>3</sup> before intervention. After January 16 the January–February three-year average increased from 53.7 μg/m<sup>3</sup> to 54.1 μg/m<sup>3</sup>, questioning the congestion charge effect.

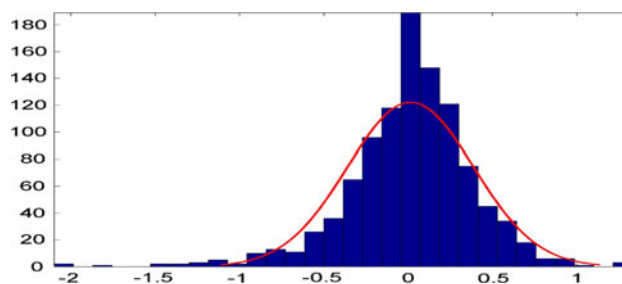
To fit the model for fine particulate matters, we use the centered log transformed concentrations which have variance  $\sigma_y^2 = 0.70$ , and, after row deletion of missing data, we get the fitted model of Table 1 based on the remaining 1038 observations. Using the seasonal structure of equation (8) for  $b$ , 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  $g = 0.212$ .

Since we are in log scale, the permanent effect, computed as both  $\hat{a}$  or  $\hat{d}$  of Eq. (1), may be interpreted as a percent change. According to this model, at Pascal station, we observe a 0.12 permanent reduction of PM<sub>2.5</sub> 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 Fig. 3 and by unreported kurtosis which is 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 were validated by simulation experiments and by comparison with robust estimation methods getting very close results to Table 1.

**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<sub>2.5</sub>. Following the approach which gives formula (7), we 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

**Table 1** Estimates and standard errors (se) of ARX model and reduction for PM<sub>2.5</sub>

Parameter	Estimate	se
$a$	0.095	0.031
PM <sub>2.5</sub> 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	



**Fig. 3** Residuals of PM<sub>2.5</sub> at Pascal station

**Table 2** Estimates and standard errors (se) of vector ARX model and reduction for PM<sub>10</sub>, single stations and city average

Station	$a$	PM <sub>10</sub> in Bergamo	Max WS winter	Average WS		$g$	$d$	$R^2$
				Summer	Winter			
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
se	0.018						0.024	

WS stands for Wind speed

$\pi = 85\%$  is given by  $m \geq m^* = 162$  as shown in Table 1, which is consistent with the above results.

### 3.2 Particulate matters

For the PM<sub>10</sub>, 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 a semi-peripheral area with pattern similar to the city center.

Table 2 shows the fitted model for PM<sub>10</sub>. We see that the persistence coefficients  $g$  are small and very close to one another, 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<sub>10</sub> 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, Fig. 4 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  $\mu\text{g}/\text{m}^3$ , so the average reduction of  $d = 0.137$  in log scale, with  $se = 0.053$ , corresponds approximately to 6  $\mu\text{g}/\text{m}^3$  for the yearly average. According to the data, this impact is stronger in Area C.

**STEM approach** In the light of the preliminary analysis, we go further with the approach of Sect. 2, obtaining the fitted model of Table 4. To avoid initial values dependence,

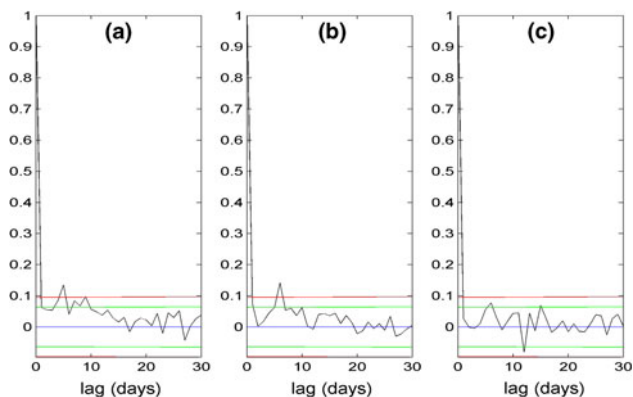
the EM algorithm has been replicated 100 times applying beta distributed random perturbations<sup>1</sup> to the initial estimates which have been computed using the method of moments. Generally speaking the standard errors are small, with the exception of the spatial correlation parameter  $\rho$ , which has a quite large uncertainty. This is not surprising since with only three stations, it is not easy to estimate spatial correlation. The reduction parameter  $\alpha$  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  $2\log(LR) = 13.2$  with a  $p$  value smaller than 0.1 %. Finally, the number of days for detection of formula (7), with nominal significance level  $p = 5\%$ , a permanent reduction of size  $\alpha^* = 0.10$  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<sub>10</sub> 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 Sect. 2.3, 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<sub>10</sub> is approximately constant around the city center but we have no information on peripheral areas.

**Nitrogen oxides** As mentioned above, NO<sub>x</sub> 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.

<sup>1</sup> 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  $E(3B) = 1$ .



**Fig. 4** Residual autocorrelations of PM<sub>10</sub> vector ARX model. **a** is for Pascal, **b** for Senato and **c** for Verziere

**Table 3** ARX model residual correlations for PM<sub>10</sub>

Città studi	1	0.65	0.58
Senato		1	0.69
Verziere			1

**Table 4** Estimates and standard errors (se) of STEM model for PM<sub>10</sub>

Parameter	Estimate	se
$\alpha$ —city center impact	0.085	0.035
PM <sub>10</sub> 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
$\gamma$	0.378	0.10
$\sigma_{\eta}^2$	0.032	0.011
$\rho$	12.4	10.1
$\sigma_{\omega}$	0.131	0.038
$\sigma_{\epsilon}^2$	0.028	0.002
Log-likelihood	3073.6	
$m^*$	154	

Using this relatively larger network, we will address the spatial variability of the congestion charge impact around the city. We first observe that NO<sub>x</sub> 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 than 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<sub>x</sub> 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 areas, at Liguria and Lambro stations, we have statistically significant increases in NO<sub>x</sub> concentrations. Note that this fact is not a model artifact but reflects the local behaviour as shown in Table 7. These points will be partially overcome in the final STEM model.

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 further 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  $t \geq t^*$ :

$$\alpha(s) = \begin{cases} \alpha_1 & s \in \text{City center} \\ \alpha_2 & s \notin \text{City center} \end{cases} \quad (10)$$

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 0.22 in log scale is estimated, and the peripheral area where NO<sub>x</sub> 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  $\rho$  is very close to the PM<sub>10</sub> case but, as expected, the standard deviation is smaller reflecting the major spatial information of nitrogen oxides network. Moreover, the number of days for detection of formula (7), with nominal significance

**Table 5** Vector ARX model for NO<sub>x</sub>

Station	<i>a</i>	Summer			Winter		NO <sub>x</sub> in Bergamo	<i>g</i>	<i>R</i> <sup>2</sup>
		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	

WS stands for wind speed, SW-PWD for south western prevailing wind direction

level  $p = 5\%$ , a permanent reduction of size  $\alpha^* = 0.20$  and power  $\pi = 85\%$  results to be  $m^* = 39$ , showing that, with this kind of spatial information, one month is sufficient to understand the permanent effect about 20% in size.

The local effects  $\bar{\omega}(s)$  of Sect. 2.3 are reported in Table 7. Note that, considering for example Pascal, if we sum up  $\alpha_1$  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.

**Table 6** Estimates and standard errors (se) of STEM model for NO<sub>x</sub>. WS stands for Wind speed

Parameter	Estimate	se
Central impact— $\alpha_1$	0.218	0.048
Peripheral impact— $\alpha_2$	−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 <sub>x</sub>	0.604	0.016
$\gamma$	0.599	0.048
$\sigma_\eta^2$	0.033	0.0047
$\rho$	12.7	3.7
$\sigma_\omega$	0.153	0.013
$\sigma_\epsilon^2$	0.054	0.002
Log-likelihood	6516.0	
$m^*$	39	



**Table 7** Estimates and standard errors (se) of STEM local effects  $\hat{\omega}(s)$  for  $\text{NO}_x$ , in log scale, and unadjusted seasonal averages before and after the intervention in *ppb*

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
Abbiategrosso	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

#### 4 Discussion and conclusions

In order to answer the three scientific questions on air quality impact raised in Sect. 1, 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 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  $\text{PM}_{2.5}$ ,  $\text{PM}_{10}$  and  $\text{NO}_x$  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  $\text{PM}_{10}$  and  $\text{NO}_x$ . 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  $3.6 \mu\text{g}/\text{m}^3$ , 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. Moreover, consistently with being mostly primary pollutants, nitrogen oxides show a larger reduction, namely 22 %, or 13 *ppb*, in the 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. As expected, in the

peripheral areas of the city, we observe only non significant changes of nitrogen oxides.

Considering the third question on spatiotemporal information, we observe that the Milan data give a substantially clear picture of air quality impact for all three contaminants considered, since they satisfy the required number of days for detection  $m^*$  for each of the three contaminants. In particular the 20 % permanent reduction of nitrogen oxides could be detected after approximately 40 days with high probability. From the spatial point of view, we observe that Milan data are sufficient for discriminating among city center and peripheral areas. Nonetheless, their spatial information content is not enough for estimating a detailed city map either for pollutant concentrations or human exposure as could be done by applying the dynamical kriging capabilities of the STEM approach proposed in this paper to a more informative monitoring network.

**Acknowledgements** This research is part of Project EN17, ‘Methods for the integration of different renewable energy sources and impact monitoring with satellite data’, funded by Lombardy Region government under ‘Frame Agreement 2009’. The help of Angela Locatelli and Francesco Miazzo for data handling has been much appreciated by the author.

#### References

- Arduino G, Fassò A (2012) Environmental regulation in the European Union. In: El-Shaarawi AH, Piegorisch WW (wds) Encyclopedia of environmetrics, 2nd edn. Wiley, New York, pp 886–889
- Bartoletti S, Loperfido N (2010) Modelling air pollution data by the skew-normal distribution. Stoch Env Res Risk Assess 24:513–517
- Box GEP, Tiao GC (1975) Intervention analysis with applications to economic and environmental problems. J Am Stat Assoc 70(349):70–79
- Bruno F, Cocchi D (2007) Recovering information from synthetic air quality indexes. Environmetrics 18:345–359
- Bruno F, Guttorp P, Sampson PD, Cocchi D (2009) A simple non-separable, non-stationary spatiotemporal model for ozone. Environ Ecol Stat 16:515–529
- Cameletti M, Ignaccolo R, Bande S (2011) Comparing spatio-temporal models for particulate matter in Piemonte. Environmetrics 22(8):985–996
- Clement L, Thas O, Vanrolleghem PA, Ottoy JP (2006) Spatio-temporal statistical models for river monitoring networks. Water Sci Technol 53(1):9–15
- Corani G (2005) Air quality prediction in Milan: feed-forward neural networks, pruned neural networks and lazy learning. Ecol Model 185:513–529
- Fassò A, Finazzi F (2011) Maximum likelihood estimation of the dynamic coregionalization model with heterotopic data. Environmetrics 22(6):735–748
- Finazzi F (2012) D-STEM: a statistical software for multivariate space-time environmental data modeling. In Goncalves AM et al (eds), Proceedings of the international workshop on spatio-temporal modelling (METMA VI). Guimaraes, 12–14 September 2012, ISBN 978-989-97939-0-3.

- Finazzi F, Fassò A (2011) Spatio temporal models with varying coefficients. In Cafarelli B (ed) *Spatial data methods for environmental and ecological processes*—2nd edn, Foggia, 1–2 September 2011. ISBN: 978-88-96025-12-3. Also in: Grasp WP, 2011, <http://www.graspa.org>, ISSN: 2037-7738.
- Finazzi F, Scott ME, Fassò A (2013) A model based framework for air quality indices and population risk evaluation. With an application to the analysis of Scottish air quality data. *J R Stat Soc C* 62(2):287–308
- Gardner MW, Dorling SR (1999) Neural network modelling and prediction of hourly NO<sub>x</sub> and NO<sub>2</sub> concentrations in urban air in London. *Atmos Environ* 33:709–719
- Gentle JE (2007) *Matrix algebra: theory, computations, and applications in statistics*. Springer, New York
- Graf-Jaccottet M, Jaunin MH (1998) Predictive models for ground ozone and nitrogen dioxide time series. *Environmetrics* 9:393–406
- Hipel KW, McLeod AI (2005) *Time series modelling of water resources and environmental systems*. © 2005 Hipel & McLeod. <http://www.stats.uwo.ca/faculty/aim/1994Book/>.
- Hong-di H, Wei-Zhen L (2012) Urban aerosol particulates on Hong Kong roadsides: size distribution and concentration levels with time. *Stoch Env Res Risk Assess* 26:177–187
- Janssen NA, Hoek G, Simic-Lawson M, Fischer P, van Bree L, ten Brink H, Keuken M, Atkinson RW, Anderson HR, Brunekreef B, Cassee FR (2011) Black carbon as an additional indicator of the adverse health effects of airborne particles compared with PM<sub>10</sub> and PM<sub>2.5</sub>. *Environ Health Perspect* 119:1691–1699
- Nadarajah S (2008) A truncated inverted beta distribution with application to air pollution data. *Stoch Env Res Risk Assess* 22:285–289
- Porcu E, Gregori P, J Mateu J (2006) Nonseparable stationary anisotropic space–time covariance functions. *Stoch Env Res Risk Assess* 21:113–122
- Soni P, Chan Y, Preissl H, Eswaran H, Wilson J, Murphy P, Lowery CL (2004) Spatial-temporal analysis of non-stationary fMEG Data. *Neurol Clin Neurophysiol* 100:1–6