"Science Stays True Here"
Journal of Mathematics and Statistical Science, Volume 2016, 781-788 | Science Signpost Publishing

Local Atmospheric Pollution Evolution through Time Series Analysis

Giuliana Passamani*, Paola Masotti

Department of Economics and Management, University of Trento, Trento, Italy.

Abstract

The monitoring and improvement of air quality are fundamental issues, given the possible effects of air pollution on human health. The analysis, in an Alpine Italian province[1], of the dynamic pattern of mutual relations among air pollutants is the main aim of the present study. In particular, the interest is the proposal of a procedure that can be used to analyse the pollution level trend. The procedure is: first, the estimation of an unobserved common component that we consider a pollution indicator and, then, the analysis of the component evolution in order to assess whether any improvement in the pollution level has been observed during the last decade. The empirical analysis is conducted through a dynamic multiple time series model with a common autoregressive stochastic factor. The results show that some improvement in the level of air pollution has been achieved, especially in the most recent years.

Keywords: air quality, air pollutants, dynamic-factor model, unobserved common factor, double-exponential smoothing.

1. Introduction

A large number of epidemiological studies have demonstrated that exposure to ambient air pollution may cause a range of human health effects mainly associated with respiratory and cardiovascular systems. These studies have pointed out that even low concentrations of air pollutants can cause adverse human health effects, WHO (2004). Furthermore, the World Health Organization (WHO) estimated that

Corresponding author: Giuliana Passamani, Department of Economics and Management, University of Trento, Trento, Italy. E-mail: giuliana.passamani@unitn.it.

[1] The daily data set for the empirical analysis has been provided by "Agenzia Provinciale per la Protezione dell' Ambiente (APPA)" of the Province of Trento (Italy).

approximately 800,000 premature deaths can be attributed to air pollution exposure in large cities (Krzyzanowski and Cohen, 2008).

To face these problems the European Union adopted a Framework Directive on Air Quality (96/62/EC) and a number of Daughter Directives regarding individual pollutants in order to assess the ambient air quality in the Member States on the basis of common methods and criteria. More recently a new Directive (2008/50/EC) on ambient air quality and cleaner air for Europe has been adopted to improve the monitoring and assessment of air quality, including pollutant emissions, and to provide information to the public. In Italy this Directive has been implemented by the Legislative Decree n° 155 of August 2010. EU legislation defines evaluation and management methods for air quality and set the standards for monitoring networks. To meet the legislation requirements, the Province of Trento has divided its territory in zones on the basis of meteorological characteristics, air emissions load and territory urbanization level. Seven fixed monitoring sites have been located in the valleys, in places corresponding to areas where the level of pollutants exceeds the upper assessment threshold established by law. Another fixed monitoring site has been located at a mountain altitude, in order to detect the air pollutants concentration in an area where pollution should be at the lowest level, with the exception of ozone which is a ubiquitous contaminant.

The Province of Trento covers an almost entirely mountain area located in the northern part of Italy, near the Austrian border. It is crossed from north to south by the river Adige valley, where the two most important cities are located: Trento, the Province capital, and Rovereto.

The aim of the present paper is to assess and better evaluate the change in ambient air quality in the Province of Trento through the estimation of a pollution indicator. The data set is made up of the time series observations relative to four pollutants: Particulate Matter (PM 10), nitrogen dioxide (NO_2), nitric oxide (NO) and ozone (O_3). The daily observations correspond to the ten-years' period starting in January 2002 and ending in December 2011. They are obtained from continuous measurements of each pollutant, the unit of measurement is $\mu g\ m^{-3}$.

For reasons due to data continuity in the observed time series, our attention is restricted to four monitoring sites: (i) the site in Trento Parco S. Chiara, which is located in a residential area of the city; (ii) the site of Rovereto Largo Posta, which is located in the city center; (iii) the site of Borgo Valsugana, a town with a population of ca. 6,000 inhabitants, in Valsugana, a valley which starts from Trento and crosses the Province from north-west to south-east; Borgo is located only few kilometers from an important highway characterized by heavy traffic levels; (iv) the site of Riva del Garda, an important touristic town located at the north-western corner of Garda Lake, characterized by wind blowing each afternoon from the lake.

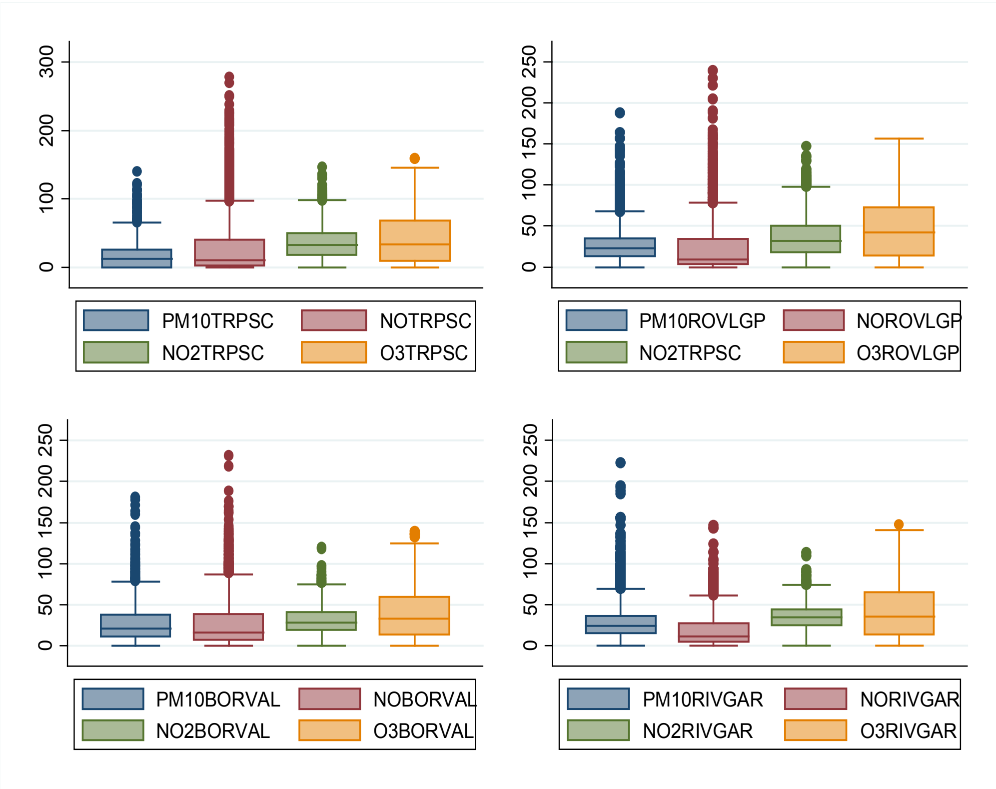

Figure 1. Box-plot of the time series for the four pollutants, PM_{10}, NO, NO_2 and O_3, recorded at the four monitoring sites during the period 01/01/2002 – 31/12/2011

Figure 1shows the box-plot representations of the four pollutants for each monitoring site. As we can notice, there are some differences in the median value of PM_{10}, NO, NO_2 and O_3 across the sites and their distribution is asymmetric with some large outliers.

The structure of the paper is the following. In Section 2 we discuss the methodological approach and the dynamic-factor model adopted for analysing the available daily time series. In Section 3 we describe the results in terms of the evolution of the estimated pollution indicator for each site and in Section 4 we conclude the paper and outline possible lines for further research.

2. The Methodological Approach

Given the aim of the paper, first we estimate the unobserved common factor. We consider the following dynamic-factor model[2]:

$$\begin{aligned} \mathbf{y}_t &= \mathbf{p} f_t + \mathbf{u}_t \\ f_t &= \alpha f_{t-1} + \nu_t, \\ \mathbf{u}_t &= \mathbf{C}\mathbf{u}_{t-1} + \boldsymbol{\varepsilon}_t \end{aligned} \qquad (1)$$

where $\mathbf{y}_t$ represents the (4×1) vector of dependent variables, the four pollutants for each site, and f_t the unobservable common factor. $\mathbf{u}_t$ and $\boldsymbol{\varepsilon}_t$ are (4×1) vectors of disturbances, ν_t is a scalar disturbance, **p** is a vector and **C** is a matrix of parameters and α is the autoregressive factor parameter. Model (1) can be considered as a dynamic factor model[3] with vector autoregressive errors. Within the statistical software Stata, it is estimated using a maximun likelihood approach implemented by writing the model in state-space form[4] and by using the stationary and the De Jong diffuse Kalman filter to derive and implement the log likelihood. In our empirical analysis, the number of autoregressive lags in the model specification and the covariance structure of the errors have been decided on the basis of some preliminary analysis, also to avoid convergence problems that can arise in the optimization phase of the likelihood, when estimating the model[5].

Once the model has been estimated, we use it in order to predict the unobserved factor variable, $\hat{f}_t$. The prediction method adopted estimates the states at each time by a Kalman smoother and using all the sample information.

Then the trend component of the predicted unobserved factor is estimated through a double exponential smoothing procedure, where the first smoothed variable, s_t, is obtained with a single exponential smoothing procedure as follows:

---

[2] Dynamic-factor models have been proposed and applied mainly in economics, where we can find a large literature originating from the work of Sargent and Sims (1977).

[3] A similar model was used by Fontanella et al. (2007) for the analysis of environmental pollution in the Milan district, following the work of Forni et al. (2000).

[4] For state-space representation the reference is Harvey (1989).

[5] Convergence problems are mentioned in Stata11 Manual.

$$s_t = \lambda \hat{f}_t + (1-\lambda)s_{t-1}, \quad 0 \le \lambda \le 1 \tag{2}$$

The double smoothed variable $s_t^{[2]}$ is the result of the same smoothing procedure applied to the smoothed variable s_t, as follows:

$$s_t^{[2]} = \lambda s_t + (1-\lambda)s_{t-1}^{[2]} \tag{3}$$

In order to get comparable trend components for the different monitoring sites, the smoothing parameter λ has been set to 0.005 in each site, which is a reasonable value leaving some seasonal behaviour in the resulting pollution level indicator, a behaviour characterising also the observed time series.

3. Results

For each monitoring site we consider $\mathbf{y}_t' = (PM_{10}, NO, NO_2, O_3)$. Assuming that these variables are stationary, we model them as linear function of an unobserved factor f_t generated by a first-order autoregressive process, and we allow the errors for the observed variables to be first-order autocorrelated. Therefore, the model is a constrained vector autoregression with an unobserved autocorrelated factor. Its ML estimation[6] has given interesting results: for each site the estimated parameters of PM_{10}, NO and NO_2 on the common factor are largely significant and with positive signs, while the estimated parameters of O_3 is significant and negative. In Figure 2 we represent the four estimated factors $\hat{f}_t$. As we can notice they are characterised by a clear seasonal behaviour.

[6] The ML estimates are not reported for reasons of space, but are available upon request.

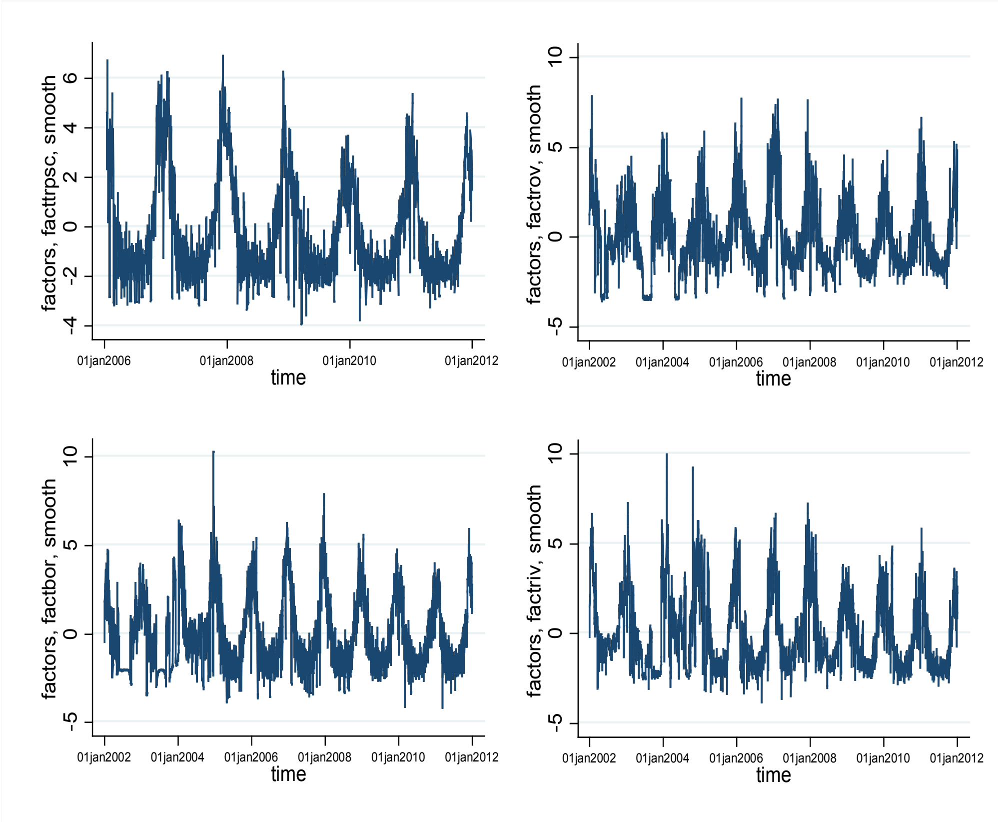

Figure 2. The estimated unobserved common factor for each monitoring site

Given $\hat{f}_t$, the four pollution indicators are obtained through the double smoothing procedure with a fixed smoothing parameter, described in (2) and (3). The representation of the smoothed common components is given in Figure 3. As we can observe the evolution of the pollution indicators show some clear different trend behaviours, though they all show an improvement in air quality in recent years. From the data relative to the monitoring site in Trento, for which the trend component is estimated from 2006 because of missing data on PM_{10}, we can see a slight worsening lasting until 2008. For the site in Rovereto we can notice higher levels of pollution in 2002 and 2003, then an improvement in 2004 and 2005, followed by higher levels in 2006, 2007 and 2008, with a peak in 2007, and then an improvement. For the site in Borgo Valsugana we can see an increasing level up to 2005, then stationary levels until 2009, when a sensible improvement has started. For the site in Riva del Garda the improvement is quite steady apart from a couple of picks in 2005 and in 2008.

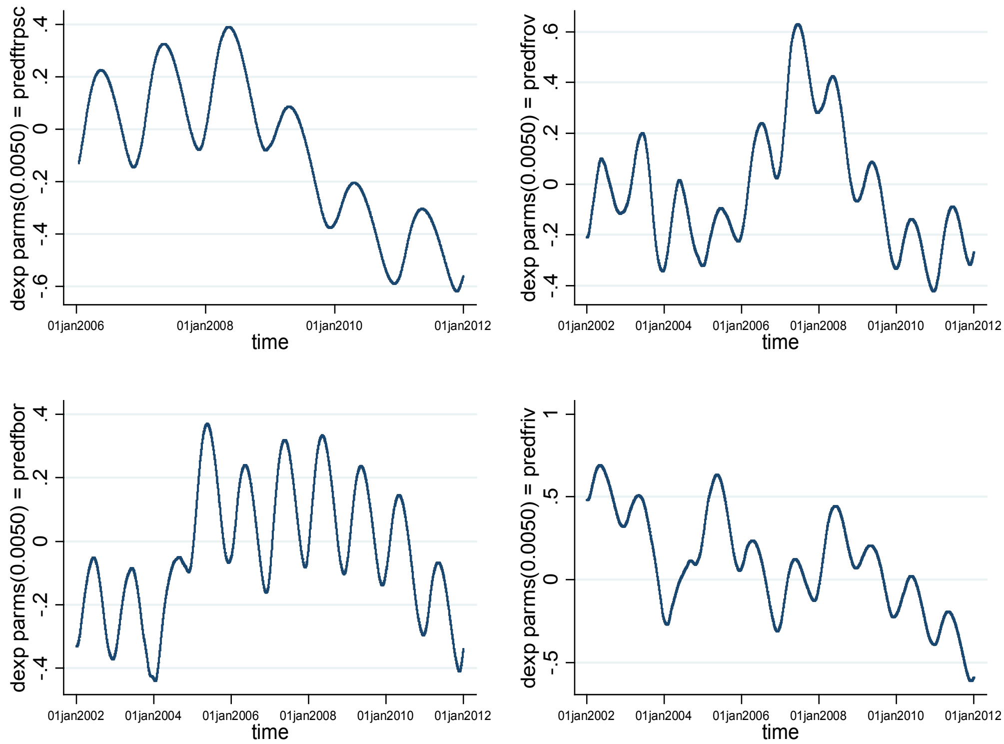

Figure 3. The estimated smoothed component of the predicted unobserved factor for the four monitoring sites

4. Conclusions

As already remarked, the aim of the paper is the proposal of a methodological procedure for estimating a pollution indicator and not to suggest a strategy for the construction of another synthetic air quality index as, among others, in Bruno and Cocchi (2007). As well as any other synthetic index, the proposed indicator summarises a complex situation in a single variable whose evolution can be compared in time and in space and the results we have obtained are interesting with respect to the aim.

In our study we haven't taken into consideration the possible effects on pollution of meteorological covariates, because we are not actually analysing how climate changes influence the level of pollution, as they are not mainly under control. We are just analysing how pollution has evolved because of best practises introduced by law. The multivariate dynamic time series model used in the analysis and tailored to the problem at hand, offers a rather simple and flexible approach to modelling an unobserved variable as pollution, given the observations on pollutants. The model could be improved by specifying a more complex specification for the covariance structure of observed variables and for the autoregressive structure of the common factor. It could be improved also by considering jointly the monitoring sites and by showing, eventually, spatial differences in the pollution indicators. In this case some identification problems should be taken into consideration, thus increasing the complexity of the analysis.

Finally, in this paper we have not considered the question of pollution forecast outside the sample. This would require a more complex time series analysis of the predicted factor, taking into consideration even its conditional heteroscedasticity as well its conditional mean.

References

[1]. Bruno, F. and Cocchi, D. (2007), Recovering Information from Synthetic Air Quality Indices, *Environmetrics*, 18, 345-59.

[2]. Fontanella, L., Ippoliti, L. and Valentini, P. (2007), Environmental Pollution Analysis by Dynamic Structural Equation Models, *Environmetrics*, 18, 265-83.

[3]. Forni, M., Hallin, M., Lippi, M., and Reichlin, L. (2000), The Generalized Dynamic-Factor Model: Identification and Estimation, *The Review of Economics and Statistics*, 82, 540-54.

[4]. Harvey, A. C. (1989) *Forecasting, Structural Time Series Models and the Kalman Filter*, Cambridge University Press, Cambridge UK.

[5]. Krzyzanowski, M., Cohen, A. (2008), Update of WHO Air Quality Guidelines, *Air Quality, Atmosphere & Health*, 1, 7-13.

[6]. Sargent, T. J. and Sims, C. A. (1977), Business Cycle Modeling without Pretending to Have Too Much a Priori Economic Theory, in Sims, C. A. (ed.) *New Methods in Business Research*, Federal Reserve Bank of Minneapolis, Minneapolis MN.

[7]. WHO (2004) *Comparative Quantification of Health Risks.Global and Regional Burden of Disease Attributable to Selected Major Risk Factors*, Ezzati, M., Lopez, A. D., Rodgers, A. and Murray, C. J. L. (eds.), World Health Organization, Geneva.

Published: Volume 2016, Issue 12 / December 25, 2016