Available online at www.sciencedirect.com
ScienceDirect
Chemosphere 69 (2007) 1055–1063
CHEMOSPHERE
www.elsevier.com/locate/chemosphere

The application of positive matrix factorization in the analysis, characterisation and detection of contaminated soils

S. Vaccaro *, E. Sobiecka, S. Contini, G. Locoro, G. Free, B.M. Gawlik

European Commission, Joint Research Centre, Institute for Environment and Sustainability, Via Enrico Fermi, 21020 Ispra, Italy

Received 25 January 2007; received in revised form 5 April 2007; accepted 10 April 2007
Available online 4 June 2007

Abstract

Multivariate factor analytical techniques are widely used for the approximation, in terms of a linear combination of factors, of multivariate experimental data. The chemical composition of soil samples are multivariate in nature and provide datasets suitable for the application of these statistical techniques. Recent developments of multivariate factor analytical techniques have led to the approach of Positive Matrix Factorization (PMF), a weighted least squares fit of a data matrix in which the weights are determined depending on the error estimates of each individual data value. This approach relies on more physically significant assumptions than methods like Principal Components Analysis which is frequently used in the analysis of soil datasets. In this paper we apply PMF to characterise the pollutant source in a set of geographically referenced soil samples taken within a 200 m radius of a site characterised by a high concentration of heavy metals. Each sample has been analysed for major and minor elements (using wavelength-dispersive X-ray fluorescence spectrometry), carbon, hydrogen and nitrogen (using a CHN elemental analyzer) and mercury (using cold-vapour atomic absorption spectrometry). Analysis of the soils using PMF resulted in a successful partitioning of variances into sources related to background soil geochemistry, organic influences and those associated with the contamination. Combining these results with a geostatistical approach successfully demarcated the main source of the combined organic and heavy metal contamination.

Keywords: Factor analysis; Positive matrix factorization; WD-XRF; CV-AAS; Elemental analysis; CHN; Soil contamination

1. Introduction

An extensive field campaign has been carried out by the European Commission's Institute for Environment and Sustainability in collaboration with the local authorities of the Pavia Province, Lombardy, Italy. The objective of the study was to perform an environmental characterisation of soils in the Pavia Province. During this study a small area identified as PS04 [see Cenci et al. (2006) for site characterisation], in a rural zone was identified as anomalous due to very high concentrations of several elements, in particular Zn, Pb, As, Cd and Hg (Cenci et al., 2006), compared to background concentrations typically found in such soil types (Utermann et al., 2006). Further sampling was carried out within a 200 m radius of this area in order to better characterise the type and extent of contamination by the analysis of several chemical parameters on geographically referenced samples (Fig. 1). Samples were taken to represent the first 30 cm soil depth. The area was of limited size having similar geology and rainfall thereby providing a relatively homogenous site for the application of multivariate factor analytical techniques. This should allow the apportionment of soil chemical composition between factors related to natural soil geochemistry and those resulting from anthropogenic influence. This should help in the identification of the source and geographic extent of the contamination.

Multivariate statistical models, like Principal Components Analysis (PCA) have yielded satisfactory results in this field of application (Lee et al., 2006; Yongming

* Corresponding author. Fax: +39 0332 78 9831.
E-mail address: stefano.vaccaro@jrc.it (S. Vaccaro).

0045-6535/$ - see front matter © 2007 Elsevier Ltd. All rights reserved.
doi:10.1016/j.chemosphere.2007.04.032

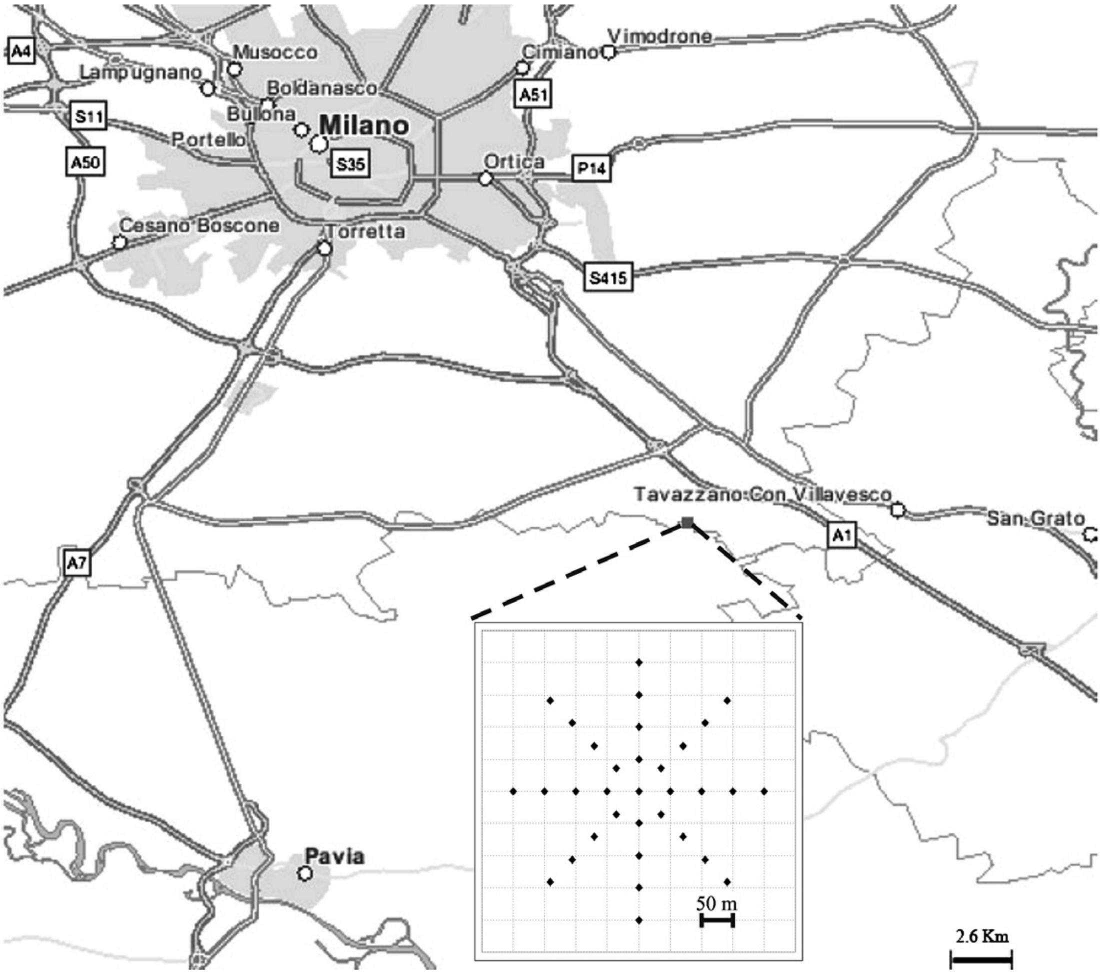

Fig. 1. Site location and sampling strategy used. The central point is the PS04 point (Map Created by the European Commission Geo-Portal – http://eugeoportal.jrc.it).

et al., 2006). However, its application is subject to limitations such as in the handling and weighting of uncertainties and noisy data (Paatero and Hopke, 2003) as well as the possible presence of negative factor loadings which are difficult to interpret in terms of positive definite physical parameters such as concentrations, masses and spectral intensities. More recently, the Positive Matrix Factorization (PMF) approach has been developed, which largely overcomes such limitations by using experimental uncertainties in the data matrix as well as by constraining the solutions to non-negative values (Paatero and Tapper, 1994). Like PCA, PMF has the advantage of being an *a posteriori* technique, in that it does not rely on previous knowledge of sources by direct measurement or from emission inventories. It has been recently applied in the examination of atmospheric pollution by particulate matter (Xie et al., 1999; Chueinta et al., 2000) and to the study of wet deposition (Anttila et al., 1995; Keeler et al., 2006). In this paper we present a new application of PMF in soil environmental chemistry using a set of geographically referenced samples. Soil samples were analysed for major elements (Si, Ca, K, Fe, Mg, Na and Al) and minor elements (Ti, S, P, Pb, Zn, Cu, Ni, Mn, Cr, V, Co, As and Cd) using wavelength-dispersive X-ray fluorescence spectrometry. Carbon, hydrogen and nitrogen were analysed using a CHN elemental analyzer and mercury using cold-vapour atomic absorption spectrometry. A simple GIS-based spatial analysis of the factors resulting from the PMF modelling was carried out using Surfer 7® software. A total of 38 points were sampled, which although limited for this kind of analysis, has allowed partitioning of geological/natural and anthropogenic factors and provides a demonstration of the potential of this approach in soil environmental geochemistry.

2. Positive matrix factorization model

2.1. Mathematical formulation

The mathematical development of PMF as a factor analytical technique is described here based on Paatero and Tapper (1993, 1994) and Paatero (1997). Analysis was carried using PMF2 version 4.2 developed by Pentti Paatero, Department of Physics, University of Helsinki.