Building and Environment 46 (2011) 577–583

Contents lists available at ScienceDirect

Building and Environment

journal homepage: www.elsevier.com/locate/buildenv

Performance assessment of air quality monitoring networks using principal component analysis and cluster analysis

Wei-Zhen Lu [a,*], Hong-Di He [a,b], Li-yun Dong [c]

[a] Department of Building and Construction, City University of Hong Kong, Hong Kong
[b] Logistics Research Center, Shanghai Maritime University, Shanghai 200135, China
[c] Shanghai Institute of Applied Mathematics and Mechanics, Shanghai University, Shanghai 200072, China

ARTICLE INFO

Article history:
Received 26 June 2010
Received in revised form
9 September 2010
Accepted 9 September 2010

Keywords:
Air quality monitoring networks
Principal component analysis
Cluster analysis
Sulphur dioxide
Respirable suspended particulates
Nitrogen dioxide

ABSTRACT

This study aims to evaluate the performance of two statistical methods, principal component analysis and cluster analysis, for the management of air quality monitoring network of Hong Kong and the reduction of associated expenses. The specific objectives include: (i) to identify city areas with similar air pollution behavior; and (ii) to locate emission sources. The statistical methods were applied to the mass concentrations of sulphur dioxide (SO_2), respirable suspended particulates (RSP) and nitrogen dioxide (NO_2), collected in monitoring network of Hong Kong from January 2001 to December 2007.

The results demonstrate that, for each pollutant, the monitoring stations are grouped into different classes based on their air pollution behaviors. The monitoring stations located in nearby area are characterized by the same specific air pollution characteristics and suggested with an effective management of air quality monitoring system. The redundant equipments should be transferred to other monitoring stations for allowing further enlargement of the monitored area. Additionally, the existence of different air pollution behaviors in the monitoring network is explained by the variability of wind directions across the region. The results imply that the air quality problem in Hong Kong is not only a local problem mainly from street-level pollutions, but also a region problem from the Pearl River Delta region.

1. Introduction

Hong Kong is one of the most developed metropolitans and characterized by the highest population and traffic densities in the world. With the continuous economic development and the population expansion, several environmental problems especially air quality problem, have become severe and attracted much attention in recent years [2,3,5,6,7,10,20]. To control and improve air quality, Hong Kong Environmental Protection Department (HKEPD) has established fourteen monitoring stations to monitor and manage. These stations are distributed in different area of Hong Kong and equipped with various instruments for detecting corresponding pollutants. Some of them are located closely in geography or characterized by similar monitoring condition, which generally present similar behaviors. It probably results in an inefficient usage of the resource and extra expense cost. Hence, it is necessary and significant to optimize the air quality monitoring network using the practical alternative methods such as principal components analysis (PCA) and cluster analysis (CA).

PCA is a statistical technique that transforms the original set of inter-correlated variables into a new set of an equal number of independent uncorrelated variables or principal components that have linear combinations to the original variables. The multicollinearity which probably implied between original variables can be removed through application of it [16,17]. On the other hand, CA is a classification method used to split a data set into a number of groups of observations which are distinct in terms of typical group values of the variables [11–13]. The aim of it is to maximize between-group variance and to minimize within-group variance.

In previous studies, these two approaches are combined to explore significant information from the origin data. Gramsch et al. [4] used the PCA and CA to determine the seasonal trends and spatial distribution of PM_{10} and O_3 in Santiago de Chile, concluding that the city had four large sectors with dissimilar air pollution behaviors. Shah and Shaheen [14] employed them to identify the major source of airborne trace metals in area of Islamabad. Pires et al. [12,13] applied them in the mass concentration of SO_2, PM_{10}, CO, NO_2 and O_3 for the efficient management of air quality in Oporto Metropolitan Area. They all verified that these two methods could complement each other and the combination of them could provide a practical alternative approach for analyzing and solving environmental problem.

* Corresponding author. Tel.: +852 27844316; fax: +852 27887612.
E-mail address: bcwzlu@cityu.edu.hk (W.-Z. Lu).

0360-1323/$ – see front matter © 2010 Elsevier Ltd. All rights reserved.
doi:10.1016/j.buildenv.2010.09.004