Environmental Technology & Innovation 5 (2016) 225–232

Contents lists available at ScienceDirect

Environmental Technology & Innovation

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

Evaluation of temporal variations in ambient air quality at Jahra using multivariate techniques

Bader Al-Anzi [a], Abdallah Abusam [b,*], Adul-Rehman Khan [c]

[a] *Environmental Technology and Management Department, College of Life Sciences, Kuwait University, P.O. Box 5969 Safat 13060, Kuwait*
[b] *WTRT Program, WRC, Kuwait Institute for Scientific Research, P.O. Box 24885 Safat, 13109, Kuwait*
[c] *EPC, Program, ELSRC, Kuwait Institute for Scientific Research, P.O. Box 24885 Safat, 13109, Kuwait*

HIGHLIGHTS

- Multivariate statistical techniques identify important pollution sources.
- Principal component analysis (PCA) discriminates sources of variations in air quality data.
- PCA reduced the number of quality parameters that need to be monitored by 40%.
- Contributions of anthropogenic pollution sources prevailed during winter seasons.

ARTICLE INFO

Article history:
Received 20 October 2014
Received in revised form 29 February 2016
Accepted 27 April 2016
Available online 2 May 2016

Keywords:
Air quality
Monitoring
Control measures
Explanatory analysis
Multivariate analysis

ABSTRACT

This study demonstrates the effectiveness of multivariate statistical methods in recognizing temporal trends and interdependency of air pollutants from large and complex air quality datasets. Eight years of ambient air quality data for the city of Jahra, Kuwait, were evaluated using various multivariate statistical techniques in order to enhance the understanding of the temporal variations in the dataset. The data are a record of 5-minute measurements of nine air quality variables (sulfur dioxide, SO_2, non-methane hydrocarbon, NM-HC, methane, CH_4, total nitrogen oxides, NO_x; as nitric oxide, NO and nitrogen dioxide, NO_2, carbon monoxide, CO, Ozone, O_3, particulate matter, PM_{10} and carbon dioxide, CO_2) and four meteorological parameters (wind speed, wind direction, ambient temperature and solar intensity). Exploratory analyses (scatter plots and box plots) and multivariate statistical analyses (principal component analysis, PCA, and correlation analysis, CA) techniques were used to assess and discriminate sources of variations in the dataset. The box plots showed a high variability in the CH_4, NM-HC and O_3 concentrations. It also showed that O_3, PM_{10}, NO, SO_2 and CO have significant seasonal patterns. CA analysis revealed significant positive correlations ($p < 0.01$) between O_3 and temperature and between PM_{10} and temperature. CA, however, also showed significant inverse correlations ($p < 0.01$) between CO_2 and temperature, and between NO and temperature. PCA allowed the identification of two different sets of 4 factors that explain 79.4% and 76.5% of the total variations in the winter and summer datasets, respectively. Furthermore, PCA resulted in a 40% reduction in the number of quality parameters. Additionally, it showed that the contributions of anthropogenic sources of air pollution (traffic, power plants and water desalination plants) prevail particularly during the winter. The obtained results are

* Corresponding author.
E-mail address: aabusam@kisr.edu.kw (A. Abusam).

http://dx.doi.org/10.1016/j.eti.2016.04.003

especially valuable for local authorities in planning analytical protocols and in designing effective air pollution control measures.

1. Introduction

Air quality assessment is of great importance because it directly influences human health and well-being and the environment. Air quality monitoring by measuring various quality parameters has increased worldwide in recent decades. Because temporal and spatial variations of air quality are often difficult to interpret, a monitoring program that provides a representative and reliable estimate of air quality has become a necessity.

Because of the high variability of air quality, intensive monitoring is usually conducted to gain reliable estimates of air quality. However, such intensive monitoring is often costly and leads to the formation of a huge and complex datasets that are comprised of a large number of physico-chemical and biological parameters. Interpretation and drawing meaningful conclusions from such datasets is usually difficult (Dixion and Chiswell, 1996). Therefore, it is necessary to optimize air quality monitoring networks and reduce the size of the generated datasets, without losing valuable information, in order to reliably express the changes in air quality.

Various statistical techniques have been used to elucidate environmental monitored data for temporal and/or spatial variation of number of pollutants. This paper proposes the use of exploratory data analysis and multivariate statistical techniques for assessing large and complex air quality datasets. Multivariate statistical techniques have proved to be appropriate tools for a meaningful data reduction and interpretation of multidimensional datasets, and therefore, they have been widely used to evaluate numerous complex environmental datasets (Singh et al., 2004; Vega et al., 1998; Noori et al., 2010). However, utilization of these statistical methods for assessing air quality problems is limited (Pires et al., 2008). The scarcity of statistical methods applied in the analysis of air quality data in the published literature sometimes reflects inconclusive trends and can be deficient in meaningful explications. The present work reflects the application of multivariate statistical tools for the interpretation of eight years of environmental filtered data obtained from Al-Jahra city in Kuwait to impart factual temporal trends and to recognize the interdependency of pollutants in order to identify common or diverse pollution sources, which is important for future planning and implementation of mitigation strategies.

2. Materials and methods

2.1. Study area

The city of Jahra is located in the northern part of Kuwait. It is inhabited by approximately 270,000 people (Alenezi et al., 2012), and it covers an area of 11,230 km^2 surrounded by several utility industries, oil fields, powers plants, and water and wastewater treatment plants. It is located on the path of a predominant northwesterly wind, which mainly blows from this direction throughout the year, thus transporting pollutants to Kuwait City.

2.2. Data

The data used in this study are from eight years (2000–2007) of air quality records obtained from the Jahra air quality monitoring station. This monitoring station is a fixed station located above the main polyclinic in the middle of a residential area. It is operated by the Kuwait Environmental Public Authority (KEPA). It measures and records the concentration of various air pollutants (CO, CO_2, CH_4, NM-HC, NO, NO_2, O_3, PM_{10}, SO_2) together with a number of meteorological parameters (wind speed, wind direction, ambient temperature and solar intensity) every 5 min. However, the data used in this study were the monthly median values of the measured air quality parameters. Median values were used instead of the mean because all the parameters were not normally distributed. Also, the ambient temperature was used as a surrogate for the meteorological conditions. The descriptive statistics of the air quality data are presented in Table 1, which clearly shows the high variability (Coefficient Variation, CV, >33%.) of the monthly concentrations of all the pollutants except CH_4 and CO_2.

2.3. Statistical analysis

The statistical analysis was performed using Matlab software version 7.0. A Box plot was used to graphically assess the monthly and seasonal variations in the collected data. In the box plot, the bottom and top horizontal lines of the box refer to the first (25%) and third (75%) quartiles, respectively, whereas the middle horizontal line (shown in red) refers to the median value. The upper and lower whisker lines show the highest and lowest data value; the plus sign refers to the outliers.

Multivariate statistics techniques were then used to reduce the large dimension of the data to extract the information that will aid in the assessment of the air pollution level. Specifically, correlation analysis (CA) and principal component