



Problems in the fingerprints based polycyclic aromatic hydrocarbons source apportionment analysis and a practical solution



Yonghong Zou ^{a, *}, Lixia Wang ^b, Erik R. Christensen ^c

^a Illinois Sustainable Technology Center, University of Illinois at Urbana-Champaign, Champaign, IL 61820, USA

^b Department of Geosciences, University of Wisconsin-Milwaukee, Milwaukee, WI 53211, USA

^c Department of Civil and Environmental Engineering, University of Wisconsin-Milwaukee, Milwaukee, WI 53201, USA

ARTICLE INFO

Article history:

Received 19 March 2015

Received in revised form

12 May 2015

Accepted 13 May 2015

Available online 23 July 2015

Keywords:

Polycyclic aromatic hydrocarbons (PAHs)

Source apportionment analysis

Fingerprints

Principal component analysis (PCA)

Positive matrix factorization (PMF)

Bayesian

ABSTRACT

This work intended to explain the challenges of the fingerprints based source apportionment method for polycyclic aromatic hydrocarbons (PAH) in the aquatic environment, and to illustrate a practical and robust solution. The PAH data detected in the sediment cores from the Illinois River provide the basis of this study. Principal component analysis (PCA) separates PAH compounds into two groups reflecting their possible airborne transport patterns; but it is not able to suggest specific sources. Not all positive matrix factorization (PMF) determined sources are distinguishable due to the variability of source fingerprints. However, they constitute useful suggestions for inputs for a Bayesian chemical mass balance (CMB) analysis. The Bayesian CMB analysis takes into account the measurement errors as well as the variations of source fingerprints, and provides a credible source apportionment. Major PAH sources for Illinois River sediments are traffic (35%), coke oven (24%), coal combustion (18%), and wood combustion (14%).

© 2015 Published by Elsevier Ltd.

1. Introduction

Source apportionment analysis is becoming a routine in investigating the occurrences and accumulations of polycyclic aromatic hydrocarbon (PAH) in natural environments, e.g. air (Harrison et al., 1996; Simcik et al., 1999), water (Motelay-Massei et al., 2007; Ross and Oros, 2004), soils (Carlon et al., 2001; Zhang et al., 2013), and sediments (Mai et al., 2003; Stout and Graan, 2010). The most frequently used source apportionment analysis methods include molecular diagnostic ratios (DRs) (Yunker et al., 2002), principal component analysis (PCA) (Jolliffe, 2002), chemical mass balance (CMB) (Li et al., 2003), and positive matrix factorization (PMF) (Paatero, 1997).

A common practice of identifying the possible PAH sources, beyond the sophisticated mathematical theories, is comparing the characteristic fingerprints preserved by different PAH sources. A fingerprint refers to the normalized composition of PAHs coherently presented in a specific source (e.g. coal, coal-tar, crude oil, soot, etc.) or an environmental matrix. The PAH fingerprints of

different sources are assumed to be unique, and are therefore often used to distinguish the origin of the PAHs. Following this idea, DR determines the possible PAH source by comparing the ratios of selected PAH compounds to the known ratios presented in specific sources (Katsoyiannis et al., 2011; Yunker et al., 2002). Similarly, CMB fits the environmental data with pre-defined (known) PAH fingerprints (Li et al., 2003). In contrast, PMF generates possible “candidate” source fingerprints, and then identifies them by comparing these to known PAH sources. However, concerns about the fingerprints based source identification are: (1) are there exclusive PAH fingerprints for different PAH sources? (2) can the fingerprints of those PAH sources remain persistent and distinguishable at the receptor? and (3) are the apportionment methods generated PAH fingerprints identical to the true sources?

Many researchers have realized the potential failure of source apportionment analysis (Galarneau, 2008; Henry, 1987; Mari et al., 2010; Robinson et al., 2006). The failure could be caused by an impaired model algorithm/theory (Henry, 1987; Zou et al., 2013) or inappropriate applications (Galarneau, 2008; Robinson et al., 2006). With respect to PAH analysis, DR has now been widely questioned for its validity in source differentiation (Dvorska et al., 2011; Galarneau, 2008; Katsoyiannis and Breivik, 2014; Katsoyiannis et al., 2011; Tobiszewski and Namiesnik, 2012). A

* Corresponding author. Illinois Sustainable Technology Center, University of Illinois at Urbana-Champaign, 1 E Hazelwood Dr., Champaign, IL 61820, USA.

E-mail address: yhzou2009@gmail.com (Y. Zou).

fundamental reason is that there is generally no clear cutoff in many of the DR values for different PAH sources (Galarneau, 2008; Katsoyiannis et al., 2011). CMB uses not just ratios between paired PAHs, but typically 7–8 compounds from low to high molecular weight which makes it more accurate and robust. However, the representative PAH source profiles are generally not present (Li et al., 2003). Inappropriate model application and interpretation is another pitfall of source apportionment analysis (Mari et al., 2010; Robinson et al., 2006).

Efforts to improve source apportionment analysis results have included physical/chemical analysis of environmental samples (Kralovec et al., 2002; Obst et al., 2011), PAH inventory history diagnostics (Bzdusek et al., 2004; Li et al., 2003), and looking for unique compound indicators (May et al., 2012). However, in the case where additional data is not available, or if further experimental analysis is not feasible, advanced data analysis techniques might provide a practical solution. Inference based computational statistical approaches, such as Monte Carlo (Sheesley et al., 2011) and Bayesian (Balachandran et al., 2013; Keats et al., 2009; Massoudieh and Kayhanian, 2013) CMB models, have the potential to overcome some of these problems and have recently attracted significant interest.

The general idea of Monte Carlo and Bayesian approaches is simulation, i.e. taking random samples from assigned or target distributions. Inferences are then made based on these random samples. The advantage of the computational statistical approaches is that they avoid the complicated numerical calculation (e.g. integration), and therefore, provide an easy solution to many scientific problems (Bolstad, 2011). The difference between Monte Carlo and Bayesian is that, in general, Monte Carlo conducts no selection during the simulation (sampling), while the Bayesian method accepts only a fraction of the randomly generated samples according to the posterior distribution. In other words, Bayesian allows the analyzer to exert more control over the model and apply his/her beliefs (prior distribution) in the simulation (Bolstad, 2011). In this study, a Bayesian based CMB method was chosen which is more easily applicable and allows us to allocate some reliance on the measurement data.

The risks of potential failure in PAH source apportionment have been reported (Galarneau, 2008; Katsoyiannis et al., 2011; Mari et al., 2010), but haven't been widely recognized. Problems in model application and interpretation have not yet been discussed in-depth. The challenge of fingerprint based source apportionment analysis is that there are often a multitude of possible sources including some that are unknown. In this study, the difficulty of comparing the similarity of PAH fingerprints is presented. We investigate, therefore, how positive matrix factorization (PMF) can be helpful in suggesting input sources. Also, an explicit demonstration is given on how the Bayesian based CMB analysis can improve the analysis. This is done using sediment PAH data from the Illinois River.

2. Methods and data

2.1. PAH fingerprints

PAHs are a large group of persistent organic pollutants (POPs) with a structure consisting of two or more fused rings (Ravindra et al., 2008). Hundreds of PAH compounds exist. However, the 16 EPA priority PAHs are the most frequently reported species (Bojes and Pope, 2007). They are acenaphthene (AcNe), acenaphthylene (AcNy), anthracene (An), benz[a]anthracene (BaA), benzo[a]pyrene (BaP), benzo[b]fluoranthene (BbF), benzo[ghi]perylene (BghiP), benzo[k]fluoranthene (BkF), chrysene (Chy), dibenz[a,h]anthracene (DBaH), fluoranthene (FlA), fluorine (Fl), indeno[1,2,3-cd]pyrene

(IP), naphthalene (Nap), phenanthrene (PhA), and pyrene (Py).

A total of 138 PAH fingerprints were collected from 25 published references (Table S1, supporting materials). These PAH sources include mineral coal, emissions from coal combustion, coke oven, coal-tar sealcoats, petroleum products (gasoline, diesel, used motor and lubricant oil), vehicle exhaust (gasoline and diesel engines), and emissions from wood combustion. Additionally, the PAH fingerprints in traffic tunnel air were also presented since they may be representative of the real on-road PAH emissions from traffic. These PAH fingerprints were considered to be able to represent the major PAH sources in the Illinois River sediments.

2.2. PAH field data

The PAH field data used in this study were generated by the Illinois River sediment reuse project named "Mud to Parks" (Marlin, 2002). More than one hundred sediment cores have been collected from the Peoria Pool of the Illinois River between 2004 and 2010. A total of 80 cores present significantly elevated total PAH concentrations and were pre-selected for analysis (Table S2). Details on sediment core sampling and preparation can be found in Cahill et al. (2008) and Slowikowski et al. (2008). After preparation, the sediment samples were sent for analysis by the contract lab, *TestAmerica Chicago*. The 16 EPA priority PAHs were analyzed using US EPA standard methods, i.e. sample processing followed EPA Method 3540C (EPA, 1996a) and chemical analysis EPA method 8270C (EPA, 1996b). Quality assurance and quality control has been carefully carried out. The surrogate recovery rates ranged from 70.8% to 135.7%.

Some PAH compounds showed concentrations below the reporting limit (RL) in some sediment cores. In order to refine the raw data, samples with more than four target PAHs (25%) below RL were not used for source apportionment analysis. Additionally, cores taken from portions of the flood plain above normal pool level, which had very short sediment deposits, were also excluded. For the retained sediment samples, the non-detects were replaced by the RL; while the measurements less than RL but higher than the method detection limit (MDL), were the approximate values suggested by the contract lab. A total of 65 sediment cores were retained for source apportionment analysis (data dimensions of 16 compounds \times 65 samples, Table S3).

2.3. Principal component analysis (PCA)

PCA is probably the oldest and best known technique of multivariate analysis (Jolliffe, 2002). Principal components (PCs) are the eigenvectors of the covariance matrix of original data. Generally, not all of these eigenvectors contribute significantly in representing the original data matrix. Instead, it is very likely that the first p ($p < m$, assuming the rank of the original data is m) eigenvectors (according to the eigenvalues from large to small) can reconstruct the original data with little loss of information. Therefore, the first p eigenvectors are selected as PCs (Zou et al., 2013).

The true meaning of the PCs is that they reflect the patterns of the correlations among the pooled elements (e.g., the m PAH compounds). Some applications view the profiles of the PCs as the fingerprints or reflecting the characteristics of the fingerprint and, furthermore, use them to identify the PAH sources. This idea makes some sense but is not guaranteed. The reason is that these PCs are a set of orthogonal variables that best span the data variances. Since they are orthogonal, it is very likely that they are not retaining the characteristics of the sources anymore (Zou et al., 2013). In general, the PCs are not identical to the fingerprints of true sources. However, the orthogonal property of the PCs makes them perfect for acting as coordinates for eigenspace projection (Zou et al., 2013).

Download English Version:

<https://daneshyari.com/en/article/6316702>

Download Persian Version:

<https://daneshyari.com/article/6316702>

[Daneshyari.com](https://daneshyari.com)