PRACE ORYGINALNE
ORIGINAL PAPERS

Medycyna Środowiskowa - Environmental Medicine 2015, Vol. 18, No. 4, 23–26
www.medycynasrodowiskowa.pl www.journal-em.com

Determination of benzo(a)pyrene content in PM10 using regression methods

Określanie zawartości benzo(a)pirenu w PM10 za pomocą metod regresyjnych

Jacek Gębicki[1] [(a,b,c)], Tomasz Ludkiewicz[2] [(d)], Jacek Namieśnik[3] [(e)]

[1] *Department of Chemical and Process Engineering, Chemical Faculty, Gdansk University of Technology, 11/12 G. Narutowicza Str., 80-233 Gdańsk, (jacek.gebicki@pg.gda.pl)*
Head of the Department: Prof. dr hab. inż. M. Kamiński
[2] *Polish Academy of Sciences, Institute of Fluid-Flow Machinery, 14 Fiszera Str.,80-231 Gdansk,*
Head of the Department: Dr hab. inż. J. Pozorski
[3] *Department of Analytical Chemistry, Chemical Faculty, Gdansk University of Technology, 11/12 G. Narutowicza Str., 80-233 Gdańsk,*
Head of the Department: Prof. dr hab. inż. J. Namieśnik

Contribution:
([a]) concept
([b]) elaboration of the text
([c]) collect literature
([d]) acquisition and analysis of data
([e]) technical supervision

Abstract

The paper presents an attempt of application of multidimensional linear regression to estimation of an empirical model describing the factors influencing on B(a)P content in suspended dust PM10 in Olsztyn and Elbląg city regions between 2010 and 2013. During this period annual average concentration of B(a)P in PM10 exceeded the admissible level 1.5-3 times. Conducted investigations confirm that the reasons of B(a)P concentration increase are low-efficiency individual home heat stations or low-temperature heat sources, which are responsible for so-called low emission during heating period. Dependences between the following quantities were analysed: concentration of PM10 dust in air, air temperature, wind velocity, air humidity. A measure of model fitting to actual B(a)P concentration in PM10 was the coefficient of determination of the model. Application of multidimensional linear regression yielded the equations characterized by high values of the coefficient of determination of the model, especially during heating season. This parameter ranged from 0.54 to 0.80 during the analyzed period.

Keywords: PM10 dust, benzo(a)pyrene B(a)P, multidimensional linear regression, coefficient of determination

Streszczenie

W pracy przedstawiono próbę zastosowania wielowymiarowej regresji liniowej do szacowania empirycznego modelu opisującego czynniki wpływające na zawartość B(a)P w pyle zawieszonym PM10 na terenie Olsztyna i Elbląga w latach 2010-2013. W tym okresie średnioroczne stężenie B(a)P w pyle PM10 ponad 1,5-3 krotnie przewyższało poziom docelowy. Przeprowadzone badania potwierdziły, że główną przyczyną wzrostu stężenia jest nieefektywność domowych kotłowni czy niskotemperaturowych źródeł ciepła, które odpowiadają za tzw. niską emisję w okresie grzewczym. Analizie poddano wzajemne zależności: stężenia pyłu PM10 w powietrzu, temperaturę powietrza, prędkość wiatru, kierunek wiatru, wilgotność powietrza. Miarą dopasowania modelu do rzeczywistego stężenia B(a)P w PM10 był współczynnik determinacji modelu. Zastosowanie wielowymiarowej regresji liniowej, przyczyniło się do opracowania równań charakteryzujących się wysokimi wartościami współczynnika determinacji modelu zwłaszcza w okresie grzewczym. Parametr ten w analizowanym okresie był na poziomie od 0,54 do 0,80.

Słowa kluczowe: pył PM10, benzo(a)piren B(a)P, wielowymiarowa regresja liniowa, współczynnik determinacji

Introduction

Since 2007 the directive no. 4 of the European Parliament and Council [1] has extended the realm of duties in the field of monitoring, evaluation and air quality management over new pollutants. One of these pollutants is benzo(a)pyrene (B(a)P), a representative of toxic and pro-carcinogenic polycyclic aromatic hydrocarbons (PAH). The directive introduced the admissible level of 1 ng/m^3 for B(a)P in suspended dust PM10 averaged over a calendar

Nadesłano: 14.08.2015
Zatwierdzono do druku: 26.10.2015

year. The main sources of B(a)P in the environment include: house heating via coal and wood combustion in furnaces and fireplaces, waste combustion, cokemaking in industrial plants, car exhaust fumes and cigarette fume [2]. The Economic Commission for Europe distinguishes several other anthropogenic sources strictly connected with energy conversion, heating, transportation and metallurgical industry. In principle, every process dealing with significant heating or incomplete combustion of organic compounds can be a source of PAH [3-5]. One of the ways PAH enter organism is their inhalation together with suspended dust particles; there is relatively strong correlation between concentration of suspended dust PM10 and PM2.5 in ambient air and concentration of B(a)P in that dust. For years the authors of literature contributions [6-9] have been analysing B(a)P content in PM10 mainly with respect to its seasonal character, when the content increases in winter season due to combustion of fuel (coal, wood) for house heating purposes and when the content of B(a)P in suspended dust decreases in summer season because of photodegradation and evaporation of B(a)P. The meteorological parameters such as temperature, humidity, solar radiation also contribute to condensation, adsorption and deposition of B(a)P on/from surface of the suspended dust [10].

The standard PN-EN 15549:2008 „Air quality. Standard method for the measurement of the concentration of benzo(a)pyrene in ambient air" describes a reference method for determination of B(a)P concentration in suspended dust. The methodology engulfs sample extraction and analysis using high-performance liquid chromatography (HPLC) with fluorescence detector (FLD). Conditioning of the filters with trapped suspended dust PM10 upon defined temperature and humidity conditions for a given period of time as well as following B(a)P analysis extend the time of B(a)P concentration measurement. Laborious and time-consuming B(a)P measurements can be avoided via estimation of B(a)P concentration with a mathematical model allowing prediction of B(a)P concentration in suspended dust with a given probability. Such approach would be especially advantageous when PM10 and meteorological parameters are measured on-line. Then the value of these would be known in real time yielding almost immediate information about potential hazards for the society. When the reference method of B(a)P measurement is used, such information is often provided after the occurrence of episodic hazard. One of the simplest methods of B(a)P concentration prediction is multidimensional linear regression based on known PM10 dust concentration and known meteorological data such as: air temperature (T), wind velocity (v), wind direction (k), air humidity (φ) and intensity of solar radiation (I) [11-13]. This method would be especially useful in the case of significant anthropogenic human activity and could become the approach supplementary to the objective methods of estimation of B(a)P concentration levels in suspended dust, which are mentioned in the Regulation of the Ministry of Environment [14]. Performing analysis of multidimensional linear regression for meteorological data and PM10 content one can obtain the following model equation (1):

$$C_{B(a)P} = a_1 C_{PM10} + a_2 T + a_3 v + a_4 k + a_5 \varphi + a_6 I + b \qquad (1)$$

where: b – free term, $a_{1\text{-}6}$ – model estimation coefficient, C_{PM10} – concentration of suspended PM10 dust, T – ambient air temperature, v – wind velocity, k – wind direction, φ – ambient air relative humidity, I – intensity of solar radiation.

A measure of model fitting to actual B(a)P content in PM10 is the model estimation coefficient.

The authors of this paper made an attempt to apply multidimensional linear regression to determine an empirical model describing the factors influencing on B(a)P content in PM10 dust in Olsztyn and Elbląg area between 2010 and 2013. The analysis engulfed interrelations between: concentration of PM10 dust in air, air temperature, wind velocity, wind direction, air relative humidity, in order to determine multi-parameter linear equations characterized by best-possible fitting to the model.

Location of measurement stations and measurement methodology

Measurement and meteorological data were obtained from two measurement stations located in Olsztyn (the first one) and Elbląg (the second one). They were typical municipal background stations and were placed beyond the range of big, point emission sources. The emission source was surface emission consisting of individual heat systems in the form of gas or oil furnaces and solid-fuel fireplaces as well as in limited extent linear emission from communication routes. The stations fulfilled the conditions of representativeness for the area characterized by given concentration level of PM10 dust.

Suspended dust PM10 was collected in the quartz filters using the reference samplers according to the PN-EN 12341standard. Dust concentration was determined from known volume of air passed through the filter and from mass of dust trapped on the filter. After weighing the filters were used for B(a)P measurement in accordance with the PN-EN 15549 standard. The methodology engulfed PAH extraction and analysis employing high-performance liquid chromatography (HPLC) with fluorescence detector (FLD). The meteorological data: air temperature, air humidity, wind velocity, wind direction and intensity of solar radiation were recorded by the sensors present in the stations.

Validation methodology of empirical model for estimation of benzo(a)pyrene concentration

Multidimensional regression consists in investigation of the influence of many independent variables on dependent variable. Such investigation can concern random variables of unknown although assumed normal distribution or the variables, the values of which originate

from time series. In both cases regression analysis is feasible. In the general case it will take the form (2):

$$y = a_1x_1 + a_2x_2 + a_3x_3 + ... + a_mx_m + b \quad (2)$$

where: $x_1...x_m$ – independent variables, $a_1...a_m$ – model estimation coefficient, b –free term.

When the equation of multidimensional linear regression is obtained, the regression coefficients are subjected to determination of statistical significance at the level of α=0.05 using the t test. After elimination of the coefficient characterized by lack of statistical significance, the process of determination of significance of the remaining coefficients of the model is repeated until all the coefficients are found statistically significant at the level of α=0.05.

A measure of model fitting of multidimensional linear regression was a coefficient of determination *D* of the model defined by the equation (3):

$$D = 1 - \frac{s^2}{s_y^2} \quad (3)$$

where: s^2 – residual variance, s^2_y – response variance of model.

Residual variance was defined by the equation (4):

$$s^2 = \frac{\sum_{i=1}^{n}(y_i - Y_i)^2}{N_\alpha} \quad (4)$$

where: y_i – independent variable (in this case concentration of B(a)P), Y_i – initial state of model according to the obtained regression, N_α – number of degrees of freedom. Response variance of model was defined by the following relation (5):

$$s_y^2 = \frac{\sum_{i=1}^{n}(y_i - \bar{y})^2}{n-1} \quad (5)$$

where: $\bar{y}$ – mean value of independent variable, *n* – number of measurements.

When the coefficient of determination of the model took the value higher than 0.6, the fitting was considered good; above 0.9 the fitting was regarded very good. Multidimensional linear regression was performed for actual data for three time periods: heating season (October-December and January-March), summer season (April-September) and entire year. Excel and STATISTICA 9 software by Statsoft Co. were the tools used for estimation of the empirical model describing the factors influencing on B(a)P content in suspended dust PM10.

Results and discussion

The calculations were performed for the 2010-2013 period using the data obtained from the municipal background stations in Olsztyn and Elbląg. Average annual concentration of B(a)P in PM10 was from 1.34 to 3.30 ng/m^3 in these stations. Table 1 contains the equations of multidimensional linear regression and coefficient of determination for different measurement periods. The presented results confirmed seasonal character of B(a)P concentration level. During the winter (heating) season the concentrations are several times higher than in the summer season. A regression model was fitted better to the data collected during the winter season. However, about 30% (1-*D*) of the total variation of B(a)P concentration depended on unknown factors, which suggests undertaking further investigations on elaboration of other mathematical models or the models, which take bigger number of independent variables into account. Wind direction parameter occurred to be statistically insignificant, whereas concentration of suspended PM10 dust, ambient air temperature, wind velocity and ambient air relative humidity were statistically significant parameters.

Application of multidimensional linear regression should become a valuable tool for evaluation of B(a)P concentration in the cases when PM10 measurements are performed and meteorological data is collected, but there is lack of information on B(a)P concentration. This method could be supplementary with respect to the objective methods of evaluation of B(a)P concentration level in suspended dust following the Regulation of the Ministry of the Environment [14]. Significant amount of data and the information from the measurement stations exposed to different levels of atmospheric pollution as well as cataloguing of emission from local sources should contribute to elaboration the model characterized by the coefficient of determination at the level of 0.9. Such approach to evaluation of B(a)P concentration seems rational as Polish energy economy is still based on natural resources and Poland is one of few countries in the European Union that exceed B(a)P admissible levels.

The attempt of analysis of B(a)P concentration level dependence on PM10 concentration, air temperature, wind velocity, air relative humidity in two measurement stations in Olsztyn and Elbląg yielded satisfactory results. This attempt was not successful only in the case of estimation of B(a)P concentration dependence on meteorological data during the summer season. The empirical model for estimation of B(a)P content in PM10 corresponds well with the model proposed by Callen [13], however it takes more independent variables (meteorological data) into account. The model proposed by Callen included such additional parameters as solar radiation, UV radiation, atmospheric pressure and rainfall, which resulted in the coefficient of determination of the model at much higher level than the one for the model proposed by the authors.

Callen and other authors also noticed that during summer season, when B(a)P concentration was at low level, it was difficult to identify unequivocally the cause of B(a)P occurrence and the proposed model of multidimensional linear regression combining B(a)P concentration with suspended PM10 dust concentration and metrological data is rather useless.

Table 1.
Coefficients of determination of multidimensional linear regression model, regression equations and average B(a)P concentration with respect to entire year, heating season and summer season. 1-Olsztyn, 2-Elbląg.

year	period	multidimensional linear regression model	coefficients of determination	concentration of B(a)P [ng/m^3]
2010	heating	1: $C_{B(a)P}$=0.18C_{PM10} -0.056φ +1.3v+0.19T+0.41 2: $C_{B(a)P}$=0.14C_{PM10}+0.028φ -0.04v-0.18T-1.96	0.54 0.71	4.94 3.15
	summer	1: $C_{B(a)P}$=0.016C_{PM10}-0.003φ -0.24v-0.04T+1.14 2: $C_{B(a)P}$=0.19C_{PM10}-0.02φ -0.07v-0.08T+2.87	0.18 0.50	0.19 1.08
	year	1: $C_{B(a)P}$=0.15C_{PM10}+0.036φ +0.72v-0.017T-5,5 2: $C_{B(a)P}$=0.13C_{PM10}-0.002φ -0.093v-0.15T+0.7	0.60 0.78	1.65 2.56
2011	heating	1: $C_{B(a)P}$=0.11C_{PM10} -0.007φ +0.4v-0.18T-0.41 2: $C_{B(a)P}$=0.19C_{PM10}+0.013φ -0.1v-0.12T-1.82	0.74 0.55	3.12 5.70
	summer	1: $C_{B(a)P}$=0.0024C_{PM10}-0.0018φ -0.009v-0.009T+0.31 2: $C_{B(a)P}$=0.0054C_{PM10}-0.0054φ +0.0012v-0.025T+0.86	0.30 0.29	0.09 0.16
	year	1: $C_{B(a)P}$=0.08C_{PM10}+0.012φ +0.14v-0.12T-0.9 2: $C_{B(a)P}$=0.17C_{PM10}+0.022φ -0.11v-0.1T-2.3	0.68 0.68	1.72 2.84
2012	heating	1: $C_{B(a)P}$=0.1C_{PM10}-0.03φ +0.27v-0.2T+3.9 2: $C_{B(a)P}$=0.2C_{PM10}+0.03φ +0.21v-0.15T-2.6	0.58 0.80	4.82 6.35
	summer	1: $C_{B(a)P}$=0.04C_{PM10}-0.02φ +0.19v-0.043T+1.26 2: $C_{B(a)P}$=0.019C_{PM10}-0.016φ -0.42v-0.11T+3.8	0.18 0.26	0.32 0.59
	year	1: $C_{B(a)P}$=0.1C_{PM10} -0.016φ +0.01v-0.19T+2.32 2: $C_{B(a)P}$=0.18C_{PM10}-0.004φ -0.08v-0.19T+1.1	0.67 0.81	2.77 3.30
2013	heating	1: $C_{B(a)P}$=0.2C_{PM10}-0.02φ +0.78v-0.17T-0.88 2: $C_{B(a)P}$=0.11C_{PM10}-0.019φ +0.18v-0.14T+1.0	0.76 0.71	3.49 3.15
	summer	1: $C_{B(a)P}$=0.04C_{PM10}-0.004φ -0.17v-0.03T+0.75 2: $C_{B(a)P}$=0.0043C_{PM10}+0.0016φ -0.15v-0.05T+1.08	0.41 0.17	0.38 0.35
	year	1: $C_{B(a)P}$=0.16C_{PM10}+0.006φ -0.19v-0.096T-0.86 2: $C_{B(a)P}$=0.07C_{PM10}-0.005φ -0.071v-0.11T+1.3	0.67 0.66	2.58 1.34

References

1. Dyrektywa Parlamentu Europejskiego i Rady nr 2004/107/WE z dnia 15 grudnia 2004 r.
2. IARC Monographs on the Evaluation of Carcinogenic Risks to Humans, vol. 92, Some Non-heterocyclic Polycyclic Aromatic Hydrocarbons and Some Related Exposures, Lyon, France, 2010.
3. Toxicological Profile for Polycyclic Hydrocarbons U.S. Department of Health and Human Services, Public Health Service, Agency for Toxic Substances and Disease Registry, Atlanta, 1995.
4. CERCLA Priority List of hazardous substances. http://www.atsdr.cdc.gov/cercla/
5. van Drooge B.L, Nikolova I., Ballesta P.P.: Thermal desorption gas chromatography-mas spectrometry as an enhanced method for the quantification of polycyclic aromatic hydrocarbons from ambient air particulate matter. J Chromatogr A 2009; 1216: 4030-4039.
6. Saarnio K., Sillanpa M., Hillamo R., et al.: Polycyclic aromatic hydrocarbons in size-segregated particulate matter from six urban sites in Europe. Atmos Environ 2008; 42: 9087-9097.
7. Callén M.S., de la Cruz M.T., López J.M., et al.: Long-Range Atmospheric Transport and Local Pollution Sources on PAH Concentrations in a South European Urban Area. Fulfilling of the European Directive. Water Air Soil Pollut 2008; 190: 271-285.
8. Amodio M., Caselli M., de Gennaro G., et al.: Particulate PAHs in two urban areas of Southern Italy: Impact of the sources, meteorological and background conditions on air quality. Environ Res 2009; 109: 812-820
9. Rehwagen M., Muller A., Massolo L., et al.: Polycyclic aromatic hydrocarbons associated with particles in ambient air from urban and industrial areas. Sci Total Environ 2005; 348:199-210.
10. van Drooge B., Ballesta P.P.: Seasonal and Daily Source Apportionment of Polycyclic Aromatic Hydrocarbon Concentrations in PM10 in a Semirural European Area. Environ Sci Technol 2009; 43: 7310-7316.
11. Akyuz M., Cabuk H.: Meteorological variations of PM2.5/PM10 concentrations and particle-associated polycyclic aromatic hydrocarbons in the atmospheric environment of Zonguldak, Turkey. J Hazard Mater 2009; 170: 13-21.
12. Lobscheid A.B., McKone T.E., Vallero D.A.: Exploring relationships between outdoor air particulateassociated polycyclic aromatic hydrocarbon and PM2.5: A case study of benzo(a)pyrene in California metropolitan regions. Atmos Environ 2007; 41: 5659-5672.
13. Callén M.S., López J.M., Mastral A.M.: Seasonal variation of benzo(a)pyrene in the Spanish airborne PM10. Multivariate linear regression model applied to estimate BaP concentrations. J Hazard Mater 2010; 180: 648-655.
14. Rozporządzenie Ministra Środowiska z dnia 13 września 2012 r. w sprawie dokonywania oceny poziomów substancji w powietrzu Dz.U. 2012 poz. 1032.

Adres do korespondencji:
Jacek Gębicki
Department of Chemical and Process Engineering, Chemical Faculty, Gdansk University of Technology, 11/12 G. Narutowicza Str., 80-233 Gdańsk
jacek.gebicki@pg.gda.pl