Developing a methodology to predict PM_{10} urban concentrations using GLM

J. M. Garcia[1], F. Teodoro[1,2], R. Cerdeira[1], L. M. R. Coelho[1] & M. G. Carvalho[3,4]
[1]*Escola Superior de Tecnologia de Setúbal, Instituto Politécnico, Portugal*
[2]*CEMAT, Instituto Superior Técnico, Portugal*
[3]*Instituto Superior Técnico, Portugal*
[4]*European Parliament, Belgium*

Abstract

A methodology using Generalized Linear Models (GLM) was developed and tested to build a model to predict PM_{10} outdoor urban concentrations. The methodology is based in the previous study of the relations between atmospheric concentrations of air pollutants CO, NO_2, NOx, VOCs, SO_2, and meteorological variables, air temperature, relative humidity and wind speed, in a particular city (Barreiro, Portugal). The model uses data from the Portuguese monitoring air quality stations network, and meteorological data. The developed GLM model consider as dependent variable PM_{10} outside air concentrations, and considers as explanatory independent variables or covariates, the air concentrations of pollutants NO_2, NOx, CO, O_3 but also the meteorological variables, air temperature, relative humidity of outside air and wind speed. A logarithmic link function was considered with a Poisson probability distribution. Particular attention was dedicated to cases with maximum air temperature below 25ºC and maximum air temperature above 25ºC. Results indicate that best performance results were achieved for model with values of maximum air temperature above 25ºC, when compared with model considering all data, or when compared with model considering maximum air temperature below 25ºC. The model was also tested with data from other Portuguese city (Oporto).
Keywords: outdoor air quality, PM10, generalized linear methods, SPSS.

WIT Transactions on Ecology and The Environment, Vol 183, © 2014 WIT Press
www.witpress.com, ISSN 1743-3541 (on-line)
doi:10.2495/AIR140051

1 Introduction

Concern about air quality has growth mainly due to the increase in respiratory problems, especially in children, elderly and people with respiratory diseases, related with air pollution. Also, the economic and social development has led to the increase of urban traffic and industry that emit a wide variety of pollutants, namely CO, NO_2, NOx, VOCs, PM and SO_2. In the last decades, traffic, air pollution and health problems related with it have increased [1]. It is now accepted that air pollutants can trigger allergies and respiratory problems, particularly in children [2–4]. In this subject PM concentration in streets is one major issue, especially in urban environments. PM contain microscopic solids or liquids, which are so small that they can penetrate deep into the lungs and cause serious health problems. In fact, in the range of air pollutants, particular attention was dedicated to Particulate Matter (PM) considering both PM_{10} and $PM_{2,5}$ [1, 5] and more recently nanoparticle [6]. Unfortunately, the knowledge of PM_{10} concentrations is not always an easy matter as in some urban locations there is no Urban Air Quality Stations Network present or even there is no way of knowing the concentrations levels. With this scenario, every way of measuring or even estimating the PM_{10} concentration of a particular site would be useful and welcome. In the past several methods have been used to estimate PM_{10} concentration levels in urban air outside locations. These methods include the monitoring of ambient concentrations with air acquisition equipment, the numerical simulation of pollutants dispersion with computers, wind tunnel experiments in urban models and statistical methods [7]. Statistical tools like statistical models based on multiple regression analysis and classification and regression trees analysis have been developed and applied in the forecasting of average daily concentrations for particulate matter and average maximum hourly ozone levels [8, 9]. In studies based on the estimation of PM concentrations using satellite remote sensing techniques, also some statistical tools have been widely used. In this field, the Aerosol Optical Thickness (AOT) is the satellite derived parameter most commonly used as the basis for PM estimation using statistics techniques [10]. Several methods have been used to correlate this satellite remote sensing (AOT) with the surface measured PM concentrations based on ground measurements from air quality stations. These include linear relations [11], statistical and chemical transport models [12], multiple regression analysis [13] and neural networks [14]. Also statistical methods were developed and used in the past to determine relationships between air pollution concentrations and meteorological parameters. Among these, methods like multiple linear regression analysis [15], nonlinear multiple regressions [16], artificial neural networks [17], and generalized additive models and fuzzy-logic-based models [18] were used. These models were tested in a perspective of daily or long term forecasting and focused in the perspective of the exploring relationship between O_3 and PM. However, in some situations it would be useful to know (or at least to estimate) unknown PM air concentration values based in values of air concentrations of other pollutants and on meteorological variables. This could be done based on known air concentrations from other air pollutants

and on meteorological parameters using data from monitoring sites or from specific acquisition data equipment. This is particularly useful in urban environments, were there is no data from monitoring sites and when it is important to know PM concentration in outside air, this is particularly important in high traffic urban sites. A well-known, documented and that tested tool like General Linear Models (GLM) [19] is used in building a methodology to estimate PM outside concentrations based on known values of outside air pollutant concentrations from the same site. GLM had shown that compared with other methods (as Multiple Linear Regression Model) can made assumptions more evident, can decouple assumptions in a better way, improving quality and showing greater flexibility. In this methodology we use data from air pollutants, CO, NOx, NO_2, O_3, SO_2 and PM_{10} that are hourly monitored by several stations, to build a model that is subsequently used to predict PM_{10} concentrations in the same site. To build this model it was taken under consideration the knowledge that Atmospheric PM are very different not only in their constitution, but also on its origin and on their govern mechanisms. Generically are grouped under the designation of particle matter (PM), a group of air pollutants considerably extended and different, and that may have their origin in sources as diverse as automobiles, steel mills, power stations, heating systems, factories cement, volcanoes, deserts and oceans, between others. In general terms, is common to consider as particulate matter the definition from NIST [20] as "any condensed-phase tridimensional discontinuity in a dispersed system may generally be considered a particle". In terms of classification, PM are usually classified based on two distinct criteria. They can be classified by their mechanism of formation, and in this case they are called primary particles or secondary particles, or can be classified by their physical size. According to the criterion of the formation mechanism, the primary particles are those that are directly emitted as particles, whereas secondary particles are those which are formed from gaseous precursors in the atmosphere through a mechanism of formation gas-particle. PM are also often classified by their physical size. Their characteristic dimensions vary from a range of few nanometres (nm) up to dozens of micrometres (μm) in diameter. The particles larger than 2.5 μm (coarse particles) are produced by mechanically breaking of the larger solid particles. This PM can include dust originating from agricultural processes transported by wind, dust originating from the bare soil, dust originating from unpaved roads or dust from other processes such as mining or stone quarrying. Smaller particles (fine particles) are mainly formed from gases. The smaller ones (less than 0.1 microns) are formed by nucleation, i.e. the condensation of substances formed by high temperature steaming or by chemical reactions in the atmosphere. The particles below 1μm may be formed by condensation of metal or condensation of organic compounds that are evaporated in combustion processes, or they can also be produced by condensation reactions resulting from atmospheric gases. The particles produced by these reactions of gases in the atmosphere are called secondary particles. Sulphate and nitrate particles are usually the predominant component of these fine particles. Other important aspect in the definition of the characteristics of PM concentrations in the atmosphere are the meteorological variables such as wind speed and

direction, atmospheric temperature, precipitation and atmospheric boundary layer height. Higher concentrations of particle concentrations are often registered under weather conditions with atmospheric stability, especially in situations of inversion with low wind speeds. Also chemical and physical processes of particle formation are regulated largely by meteorological variables [21]. Chaloulakou *et al*. [22] found that $PM_{2.5}$ and PM_{10} concentrations were highly correlated with carbon monoxide, black carbon and nitrogen oxides and inversely correlated with local wind speed. Also solar radiation and temperature have major importance in the mechanisms of formation of secondary particles. Results from Anderson indicate 25ºC as key air temperature value from which the occurrence of summertime air pollution episodes are promoted [23].

The purpose of this paper is study the relationship between atmospheric pollutants and develop a methodology that estimate PM_{10} concentrations in the city of Barreiro in Portugal, by using an Generalised Linear Model (GLM), using data from air quality stations with measured concentration values of CO, NOx, VOCs, and SO_2 to predict the values of PM_{10} concentration. The predict values are compared with real measured values of PM_{10} outside air concentrations in the city.

2 Methods

2.1 Location

Barreiro is a medium size city located 40km south of Lisbon, Portugal, with 34km^2 area and about 80000 inhabitants, with industry near the centre and typical suburbs important car traffic fluxes. The city is almost flat, with highest point at approximately 10 meters above sea level. The weather is temperate, with no severe seasons. The main industrial activity in Barreiro city is developed in the industrial area. A natural gas power plant and some chemical industries are the main industrial sources. The most important pollutants released from these industrial sources are NOx, SO_2 and PM.

2.2 Meteorological and air quality data

Meteorological data was provided by the Instituto Português do Mar e Atmosfera (IPMA). The prevailing wind direction is NW (frequency 35.1%). The highest wind speed registered correspond to the prevailing direction NW (14,1 km/h). The NW wind is particularly frequent in the summer months (June, July and August), with a maximum occurring in August (58.5%) and a minimum frequency recorded in December (15.6%). The average wind speed is relatively constant throughout the year. Air Quality data from pollutants concentrations (CO, NOx, NO, NO_2, O_3, SO_2 and PM_{10}) are hourly monitored by seven air quality stations that are managed by the Portuguese government. Data from September 2003 to December 2005 was statistically treated, according to the pollutant in question. A twenty-four hour mean was calculated for NOx, NO,

NO_2, SO_2 and PM_{10} and 8 hours mean to CO and O_3. Daily averages of each pollutant were related with each other and with meteorological data.

2.3 The GLM methodology

A General Linear Models (GLM) [19] was used to building a methodology to estimate PM outside concentrations based on known values of other outside air pollutant concentrations. GLM are based in the assumption that there are K independent values Y_1, ..., Y_K, from a variable of interest or response variable (effect) that follows an exponential family distribution with expected value E (Yi) = μi [24]. Considering K vectors xi = (1 xi1 x12 … xip)^t, i=1, ..., K, containing the values of p explanatory variables, independent or covariates (variables candidate to "causes"). Considered also a link differentiable function g, such that:

$$g(\mu_i) = x_i^t\beta,\ i = 1, \ldots, K \tag{1}$$

where (β = β1 β2 … βp) are the values of parameters to be estimated. Thus if we consider for the function g the identity function we have:

$$g(\mu_i) = \mu_i \tag{2}$$

then

$$\mu_i = E(Y_i) = x_i^t\beta \tag{3}$$

The resulting model is the Gaussian linear regression model. If alternatively, consider the function g as a logarithmic function and Yi has a Poisson distribution, then the model will result in a Poisson regression model and each term β_i is the effect of variable X_i in g (μi). Each β_i represents the "effect" of variable Xi in the function g(μi).

In this case the objective is to estimate PM_{10} concentration values based on other variables, like air pollutant concentration from CO, NO_2, NO_x, O_3 and SO_2 (in μg/m^3) and meteorological variable as air temperature (T,ºC), relative humidity (RH,%) and wind velocity (WV, m/s). The general model parameter used in GLM models are resumed in table 1. Statistical Package software for

Table 1: General model parameter information resume.

Dependent variable	PM_{10} concentration (μg/m^3)
Covariates	CO concentration (μg/m^3) NO_2 concentration (μg/m^3) NO_X concentration (μg/m^3) O_3 concentration (μg/m^3) SO_2 concentration (μg/m^3) Temp (ºC) RH (%) WV (m/s)
Probability distribution	Poisson
Link function	Logarithmic

Social Sciences SPSS 10.0 for windows was used to build and analyse the model.

3 Results

GLM models were used to investigate the complex relationships between the concentration of 5 air pollutant concentrations, meteorological and PM_{10} concentration levels in the Barreiro city.

$$\ln[PM_{10}] = \alpha + \beta_1 var_1 + \beta_2 var_2 + \beta_3 var_3 + \cdots + \beta_i var_i \quad (4)$$

Based on these results, particle concentrations estimations for PM_{10} can be expressed as the product of the exponential terms:

$$[PM_{10}] = e^{(\alpha+\beta_1 var_1+\beta_2 var_2+\beta_3 var_3+\cdots+\beta_i var_i)} \quad (5)$$

The first term contains the regression intercept and the rest terms contain binary variables, originated from GLM model as explained above. This methodology as applied to three tested models A, B, and C. The three models presented in table. 3 differ only in data considered. In model A, we considered the total number of observations recorded. In model B we considered the observations recorded in days with maximum air temperature of day above 25ºC (maximum). In model C we consider only observations with maximum air temperature of day less or equal to 25ºC. These considerations are shortly resumed in table 2.

Table 2: Specific models short description.

Model	Restriction	Dependent Variable	Covariates
A	No - All values	PM_{10} conc. ($\mu g/m^3$)	CO, NO_2, NO_x, O_3 SO_2 T, RH, WV
B	Tmax > 25ºC	PM_{10} conc. ($\mu g/m^3$)	CO, NO_2, NO_x, O_3 SO_2 T, RH, WV
C	Tmax ≤ 25ºC	PM_{10} conc. ($\mu g/m^3$)	CO, NO_2, NO_x, O_3 SO_2 T, RH, WV

The β coefficients obtained with methodology implemented for the three models are:

Model A:
Ln PM_{10} = 2,425652 - 0,000357 [CO] + 0,001821 [O_3] - 0,000364 [SO_2] + 0,028348 [NO_2] + 0,000093 [NOx] + 0,016820 Temp - 0,000490 HR+ 0,002821WV (6)

Model B:
Ln PM_{10} = 1,957605 - 0,000204 x [CO] + 0,001931 [O_3] - 0,003097 [SO_2] + 0,024388 [NO_2]+ 0,000309 [NOx]+ 0,043356 Temp - 0,000960 HR + 0,003548WV (7)

Model C:
Ln PM_{10} = 2,419685 - 0,000219 [CO] + 0,000863 [O_3] + 0,002149 [SO_2] + 0,019767 [NO_2] + 0,001449 [NO_x] + 0,021912 Temp + 0,000153 HR + 0,003008WV (8)

Fig. 1 shows the scattered plot with measured PM_{10} concentrations versus the PM_{10} concentration values predicted by the three models (A, B and C). The data

values of measured PM_{10} concentrations are identified by PM_{10} measured in $\mu g/m^3$ and the PM_{10} concentrations values predicted by the three models are identified by predicted mean of response in $\mu g/m^3$.

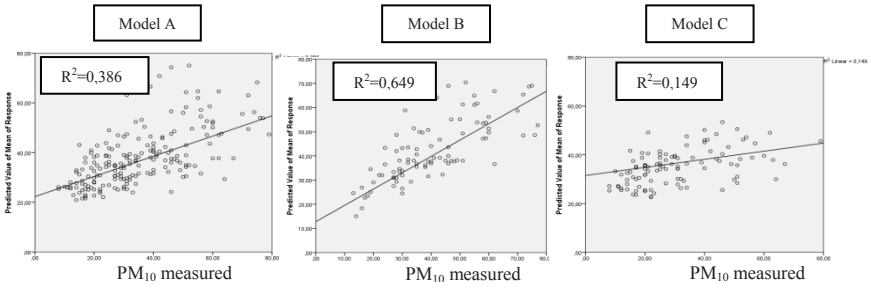

Figure 1: Comparison between PM concentrations predicted versus measured for the three models.

By the analysis of Fig. 1 and knowing that the correlation coefficient R^2x100 gives the percentage of variability explained by the model (R^2 = sum of squared residuals (SSR)/total sum of squares (SST)), the calculation of R2 results that model B is the best of the three selected models, with $R^2 = 65\%$. Models A and C have a weak explanatory capacity. Table 3 shows a resume of the statistical model results performance for the three models (A, B and C).

Table 3: Resume of models results performance.

	Model A			Model B			Model C		
	Value	df	Value/df	Value	df	Value/df	Value	df	Value/df
Deviance	26868,641	6718	4,000	6450,747	3263	1,977	17576,459	3220	5,459
Scaled Deviance	26868,641	6718		6450,747	3263		17576,459	3220	
Pearson Chi-Square	27483,708	6718	4,091	6471,573	3263	1,983	18044,189	3220	5,604
Scaled Pearson Chi-Square	27483,708	6718		6471,573	3263		18044,189	3220	
Log Likelihood[b]	-31497,004			-12102,567			-17428,695		
Akaike's Information Criterion (AIC)	63012,008			24223,134			34875,391		
Finite Sample Corrected AIC (AICC)	63012,034			24223,189			34875,447		
Bayesian Information Criterion (BIC)	63073,333			24277,973			34930,11		
Consistent AIC (CAIC)	63082,333			24286,973			34939,11		

The first column of Table 3 presents the statistics tests most often used in generalized linear models and represent measures of dispersion (generalized and/or corrected), which permit to test the quality of models. Values from Table 3, confirm that model B is the one with best performance results shown by statistical tests. These statistics tests are obtained using all the deviations obtained between the estimated and recorded (residuals) for each observation. Considering the Akaike Information Criterion, the objective is to minimize AIC. From the three models, model B is the one with lowest AIC, which means that evidence for the model B is the best. The same can be concluded when analysing

AICC (Akaike Information Criterion corrected by minimizing the number of model parameters). When comparing with the quantile of a chi-square distribution with n-p degrees of freedom (n-number of observations, p-number of estimated parameters) it is possible to measure the suitability of models Results of deviance show that the three are suitable. Another measure of goodness of fit is the Pearson chi-square test, which leads to the same conclusions when compared with the quantile of the chi-square distribution with n-p degrees of freedom. Table 4 shows the likelihood ratio chi-square test, which compares each model with the null model. Regardless of model B is considered the best, each model individually, has a greater explanation of the dependent variable using some of the explanatory than any other model without explanatory variables.

Table 4: Models likelihood ratio chi-square test performance.

	Model A			Model B			Model C		
Omnibus Test	Value	df	Value/df	Value	df	Value/df	Value	df	Value/df
Likelihood Ratio Chi-Square	17979,211	8	0,000	14532,589	8	0	3537,432	8	0

In Fig. 2, we observe that the residues associated with the model B are those with a more adequate to the expected aspect: cloud without standard and with homogeneous variability (white noise). Either model A or model C, the residues appear to have a functional relationship and not look like white noise. The variability is also not constant as would be expected. Some diagnostic tests have been made (independence, heteroscedescidade, normality) and models A and C are rejected. Only after validation of residuals has behavior of white noise with normal distribution is that it can and should consider the inference using models.

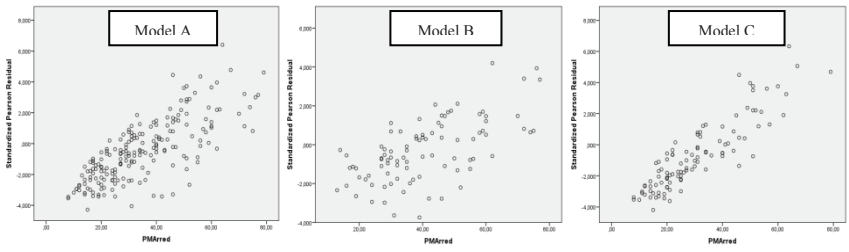

Figure 2: Scattered plot of residuals for the three models.

One last step for evaluating the quality of the model is to perform simple tests using the Wald Chi-Square statistic (Table 5). This test serves to verify that some independent variable (explanatory) in particular, contributes significantly to the explanation of the response variable, testing in the form H0: $\beta_i = 0$ versus H1: $\beta_i \neq 0$. If we reject the null hypothesis, we have evidence that the variable is a good explanatory variable. From Table 5, the p-values (sig in Table 5) associated with the nullity test of each parameter, the sig values are zero in majority,

indicating rejection of the null hypothesis, showing that the associated variables should be considered. Note that the model HR variable B is statistically significant (p-value = 0.021), not being in model A (p-value = 0:08) and model C (p-value = 0.72). Remember that we rejects the null hypothesis if p-value <significance level (the level of significance is usually 5%). It is concluded that the relative humidity is important when considering the higher temperatures.

Table 5: Models hypotheses tests.

Parameter	Model A				Model B				Model C			
	B	Wald Chi-Square	df	Sig.	B	Wald Chi-Square	df	Sig.	B	Wald Chi-Square	df	Sig.
(Intercept)	2,425652	7677,888	1	0,000	1,957605	1146,149	1	0,000	2,419685	2184,525	1	0,000
CO	-0,000357	643,828	1	0,000	-0,000204	109,603	1	0,000	-0,000219	84,91	1	0,000
O3	0,001821	195,808	1	0,000	0,001931	146,327	1	0,000	0,000863	7,1	1	0,008
SO2	0,000364	22,99	1	0,000	-0,003097	359,387	1	0,000	0,002149	530,754	1	0,000
NO2	0,028348	8976,73	1	0,000	0,024388	3932,244	1	0,000	0,019767	785,657	1	0,000
NOx	0,000093	10,957	1	0,001	0,000309	83,629	1	0,000	0,001449	200,609	1	0,000
Temp	0,01682	613,923	1	0,000	0,043356	686,462	1	0,000	0,021912	250,801	1	0,000
HR	-0,00049	3,059	1	0,080	-0,00096	5,355	1	0,021	0,000153	0,129	1	0,720
Wind	0,002821	1136,725	1	0,000	0,003548	167,289	1	0,000	0,003008	979,439	1	0,000

It is important to realize the significance of βi coefficients associated with each explanatory variable vari from model. If the variable vari changes Δvari units, keeping the remaining explanatory constants, we have

$$\Delta(\ln PM_i) = \Delta(\beta_i var_i) \Leftrightarrow \frac{\Delta PM_i}{PM_i} = \beta_i \Delta var_i \, . \quad (9)$$

4 Model implementation to Oporto data

Knowing that model B (Tmax air > 25ºC) is the model that best predicts PM10 concentrations based in measured concentrations of CO, NO_2, NOx, O_3 SO_2 T, RH, WV, model B was attested with data from a different Portuguese city (Oporto) following the same methodology and the same coefficients. Data from the Portuguese air quality network managed by CCDR, from Campanhã air quality station was used, considering values from January 2011 to December 2011. Also meteorological data from FEUP meteorological acquisition station, at the same period was used. Results showing PM_{10} concentrations predicted by the model and PM_{10} concentrations measured are shown in Fig. 3.

Inspection of fig. 3 shows that the model predicts with reasonable accuracy PM_{10} concentrations in Oporto (R^2=0.4705). It is also visible that the model predict poorly PM_{10} concentrations in the range of <20μg/m^3.

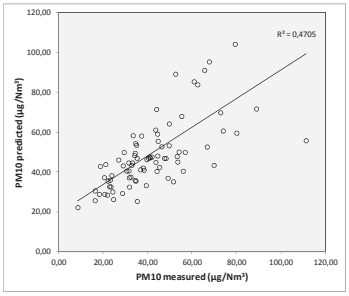

Figure 3: Comparison between PM concentrations predicted versus measured for Oporto data.

5 Conclusions

Results analysis show that model A (all values) predict very bad results for PM_{10} concentrations (R^2=0,386) with worst results for concentration values below 20 ($\mu g/m^3$). With the objective of improving the model accuracy two sub models with the criteria of maximum temperature of air above 25ºC (model B) and below 25ºC (model C) were developed.

By the knowledge of PM secondary formation mechanisms discussed in section it is expected that PM concentration could be correlated with gaseous pollutants mainly NOx SO_2, VOC and temperature. For this specific case we have no VOC data so VOC concentrations were not used. For O_3 knowing that this pollutant is also a result of photochemical oxidation is expected that O_3 could also been correlate with secondary PM even if O_3 is not a precursor for secondary particles. Comparisons of the three models show that best performance results are achieved for model B that considers only data with values of Tmax air above 25ºC (R^2=0,649) in accordance with results from (Anderson *et al*. [23]) which concludes that "simultaneous occurrence of daily maximum temperatures above 25˚C and low wind speed conditions which favour the occurrence of summertime air pollution episodes". When comparing model A (all data) and model B (Tmax air > 25ºC) and model C (Tmax air < 25ºC) the best fit prediction is achieved in model B showing the importance of higher air temperature in the formation of the secondary particles in air. This can also be concluded by observation of the relatively high coefficient values in Temperature variable observed.

Results show a good accuracy for situations were solar radiation is an important factor, which is reflected in the outsider air temperature parameter Tmax air >25ºC. These model are an important tool in situations where there is no measurements of PM concentrations but it is possible to achieve data from other gaseous air pollutants as CO, NO_2, NOx, O_3 SO_2 and also meteorological data as T, RH and WV.

Acknowledgements

The authors wish to acknowledge Comissão de Coordenação e Desenvolvimento Regional de Lisboa e Vale do Tejo (CCDR-LVT) and Instituto Português do Mar e Atmosfera (IPMA) for the information provided.

References

[1] WHO, (2004). World Health Organization, "Outdoor air pollution: assessing the environment burden of disease at national and local levels".
[2] E.E.A., (2005), "Environment and Health". EEA report, Denmark – Copenhagen, No 10/2005.
[3] EPA (2002). "Child-specific exposure factors handbook". U.S. Environmental Protection Agency.
[4] Pekkanen J, Remes ST, Husman T, Lindberg M, Kajosaari M, Koivikko A, Soininen L, (1997). "Prevalence of asthma symptoms in video and written questionnaires among children in four regions of Finland", Eur Respir J; 10: 1787-1794.
[5] EEA, (2011) European Environment Agency "*Air quality in Europe — 2011 report*", European Environment Agency.
[6] Kumar, P., Ketzel M., Vardoulakis S., Pirjola L., Britter R., (2011), "*Dynamics and dispersion modelling of nanoparticles from road traffic in the urban atmospheric environment – a review*", Journal of Aerosol Science, 42, 580-603.
[7] Gilbert, R. O. (1987). "Statistical Methods for Environmental Pollution Monitoring". John Wiley & Sons. ISBN 0-471-28878-0.
[8] Neto, J., Torres P., Ferreira F. Boavida F., (2009). "Lisbon air quality forecast using statistical methods". Int. J of Environmental Pollution. Vol. 39, N3/4, 333-339.
[9] Demuzere M., van Lipzig N. P. M. (2010). "A new method to estimate air-quality levels using a synoptic-regression approach. Part I: Present-day O_3 and PM_{10} analysis". Atmospheric Environment 44, 1341-1355.
[10] Themistocleous, K., Hadjimitsis, D.G., Retalis, A., Chrysoulakis, N., (2012). "The development of air quality indices through image-retrieved AOT and PM10 measurements in Limassol Cyprus". Proceedings of SPIE Remote Sensing 2012: Remote Sensing of Clouds and the Atmosphere XVII and Lidar Technologies, Techniques, and Measurements for Atmospheric Remote Sensing VIII. Proc. of SPIE, vol. 8534, p. 85340B. http://dx.doi.org/10.1117/12.974701.
[11] Yap, X.Q., Hashim, M., (2012). "A robust calibration approach for PM10 prediction from MODIS aerosol optical depth". Atmospheric Chemistry and Physics Discussions 12, 31483-31505. http://dx.doi.org /10.5194/acpd-12-31483-2012.
[12] Kloog, I., Nordio, F., Coull, B.A., Schwartz, J., (2012). "Incorporating local land use regression and satellite aerosol optical depth in a hybrid model of spatiotemporal PM2.5 exposures in the Mid-Atlantic States".