MULTIVARIATE ANALYSIS OF RESPIRATORY DISEASES AND THEIR ASSOCIATION WITH METEOROLOGICAL PARAMETERS AS WELL AS BIOLOGICAL AND CHEMICAL AIR POLLUTANTS

László MAKRA[1], Ioana IONEL[2], Zoltán CSÉPE[1], István MATYASOVSZKY[3], Nicolae LONTIS[2], Francisc POPESCU[2], Zoltán SÜMEGHY[1]

[1]University of Szeged, Hungary, [2]„Politehnica“ University, Timişoara,
[3]Eötvös Loránd University, Budapest, Hungary

REZUMAT. Scopul studiului este de a analiza efectul comun al poluanților atmosferici biologici (polen) şi chimici, precum şi variabilele meteorologice cu privire la internările pacienților cu probleme respiratorii, pentru diferite grupe de vârstă în diferite anotimpuri, din regiunea Szeget din sudul Ungariei. Setul de date folosit este unic, în sensul că include toate cele trei categorii. Variabilele meteorologice şi poluanții chimici atmosferici s-au colectat în stația de monitorizare situată în centrul oraşului Szeget.

Cuvinte cheie: poluant atmosferic, polen, date meteorologice, pacient cu probleme respiratorie.

ABSTRACT. The aim of this study is to analyse the joint effect of biological (pollen) and chemical air pollutants, as well as meteorological variables on the hospital admissions of respiratory-problem patients of different age groups during different seasons in the Szeged region in Southern Hungary. The data set applied is unique in the sense that includes all of the above three categories of influencing variables. Meteorological variables and chemical air pollutants were collected in a monitoring station located in the inner city area of Szeged.

Keywords: air pollutant, pollen, meteorological data, respiratory problem patient.

1. INTRODUCTION

Air pollution, as a major and constantly growing risk for the environment, is associated with a large rise in medical costs and is estimated to cause about 800,000 premature deaths annually worldwide (Cohen et al., 2005). The prevalence of allergic respiratory problems has also increased during the past three decades, especially in industrialised countries. This increase may be explained by changes in environmental factors (D'Amato et al., 2005). Weather conditions can also affect both the biological and chemical air pollutants. There is also evidence that air pollution increases exposure to the allergens, their concentration and/or biological allergenic activity (Just et al., 2007).

Air pollution in Hungary is one of the highest in Europe. Around 16,000 annual premature deaths attributable to exposure to ambient PM_{10} concentrations are estimated in the country (Ågren, 2010). Furthermore, airborne pollen levels are also high. The Carpathian basin, including Hungary, is considered the most polluted region with airborne ragweed *(Ambrosia)* pollen in Europe (Makra et al., 2010). In Szeged, 83.7% of patients suffering from respiratory problems are sensitive to *Ambrosia* pollen (Makra et al., 2010).

The purpose of this study is to analyse the joint effect of biological (pollen) and chemical air pollutants, as well as meteorological variables on the hospital admissions of respiratory-problem patients of different age groups during different seasons in the Szeged region in Southern Hungary. The data set applied is unique in the sense that includes all of the above three categories of influencing variables. This study analyses one of the largest data sets used in the literature on respiratory hospital admissions.

2. MATERIALS AND METHODS

2.1. Location and data

Szeged (46.25°N; 20.10°E) is the largest settlement in South-eastern Hungary. The city is the centre of the Szeged region with 203,000 inhabitants.

Meteorological variables and chemical air pollutants were collected in a monitoring station located in the inner city area of Szeged. Daily values taken were for the mean temperature (*T*, °C), mean global solar flux (*GSF*, $W \cdot m^{-2}$), mean relative humidity (*RH*, %), mean sea level air pressure (*P*, hPa) and mean wind speed (WS, $m \cdot s^{-1}$). Chemical air pollutants include the daily average mass concentrations of CO, NO, NO_2, SO_2, O_3 and PM_{10} ($\mu g \cdot m^{-3}$) (Alves et al., 2010). Two pollen variables were formed for our analysis: the pollen level of *Ambrosia*, and the total pollen count (the pollen counts of each of the 24 taxa examined) excluding the pollen of *Ambrosia*. Both pollen variables were considered for the pollen season of *Ambrosia* (July 15 – October 16).

The daily number of hospital admissions recorded with respiratory problems comes from the Thorax Surgery Hospital in Deszk, located about 10 km from the monitoring station in Szeged. Due to the very small number of younger patients (0-14 years) only three groups, namely adult patients (15-64 years), elderly patients (65 years of age or above), as well as all patients including the younger age group were analysed. The population consists of 133,464 hospital admissions of subjects resident in Szeged (Table 1).

The analysis was performed for the nine-year period 1999-2007 with two data sets according to the pollen season of *Ambrosia* (July 15 – October 16) and to the pollen-free season (October 17 – January 13). Note that Saturdays, Sundays and holidays as days without hospitalisation were excluded from the analysis. The pollen season defined by Galán et al. (2001) varies from year to year; here the longest observed pollen season during the nine-year period was considered for each year.

Table 1. **Parameters of daily respiratory admissions based on different age categories and seasons**

Parameter	Age categories		
	15-64 years	**Over 65 years**	**All age groups**
	Pollen season of *Ambrosia*		
Total number	81,348	13,776	95,251
Mean	83.10	11.75	95.01
Standard deviation	36.01	5.65	39.67
	Pollen-free season		
Total number	31,686	6,474	38,213
Mean	59.34	12.12	71.56
Standard deviation	23.29	6.32	27.73

2.2. Methods

2.2.1. Cluster analysis

Cluster analysis is a common statistical technique for objectively grouping elements. The aim is to maximise the homogeneity of elements within the clusters and to maximise the heterogeneity among the clusters. Here a non-hierarchical cluster analysis with k-means algorithm using a Mahalanobis metric (Mahalanobis, 1936) was carried out. The data to be clustered include daily values of the 13 explanatory variables (5 meteorological elements, 6 chemical pollutants and 2 pollen types). The homogeneity within clusters was measured by RMSD defined as the sum of the root mean square deviations of cluster elements from the corresponding cluster centre over clusters. The RMSD value usually decreases with an increasing number of clusters. Thus, this quantity itself is not very useful for deciding the optimal number of clusters. However, the change of RMSD (CRMSD) or even the change of CMRSD (CCRMSD) versus the change of cluster numbers is much more informative (Makra et al., 2010).

2.2.2. Analysis of variance (ANOVA)

A one-way analysis of variance (ANOVA) is used to determine whether the inter group variance is significantly higher than the intra group variance of a data set. After performing ANOVA on the averages of the groups in question, a post-hoc Tukey test is applied to establish which groups differed significantly from each other (Tukey, 1985). Significant differences among mean hospital admissions corresponding to different cluster pairs may reveal an important influence of the meteorological elements, chemical air pollutants and given pollen types on the daily number of respiratory admissions.

2.2.3. Factor analysis and special transformation

Factor analysis identifies linear relationships among subsets of examined variables and this helps to reduce the dimensionality of the initial database without substantial loss of information. First, a factor analysis was applied to the initial dataset consisting of 14 variables (13 explanatory variables and 1 resultant variable defined by the number of daily hospital admissions with respiratory problems) in order to transform the original variables to fewer variables. These new variables (called factors) can be viewed as latent variables explaining the joint behaviour of weather-pollutant-hospital admission variables. The optimum number of retained factors is determined by different statistical criteria (Jolliffe, 1993). The most common and widely accepted one is to specify a least percentage (80%) of the total variance in the original variables that has to be achieved (Liu, 2009). After performing the factor analysis, a special transformation of the retained factors was made to discover to what degree the above-mentioned explanatory variables affect the resultant variable, and to give a rank of their influence (Jahn and Vahle 1968).

2.2.4. Time-varying multivariate linear regression with time lags

The task is to establish a relationship between explanatory variables and the resultant variable. As both kinds of variables exhibit annual trends, regression coefficients in the linear relationship have annual courses described by sine and cosine functions with yearly and half-yearly periods. This latter cycle was introduced to describe the asymmetries of the annual courses. The coefficients of these periodic functions were estimated using the least squares principle.

It is reasonable to allow time lags between pollutants concentrations and the number of hospital admissions. Therefore, the univariate version of the above-mentioned time-varying linear regression was carried out with each individual explanatory variable with different time lags including a zero lag. The time lags that minimise the mean square errors were regarded as optimal.

3. RESULTS

3.1. Cluster analysis and ANOVA

A cluster analysis for the pollen season of *Ambrosia* and the pollen-free season resulted in five and four clusters, respectively (Tables 2a-b).

The analysis of variance revealed a significant difference at least at a 95% probability level in the mean values of patient numbers among the individual clusters. The Tukey test indicated significant differences both for adults, the elderly and all age groups among the mean patient numbers of the cluster pairs. Only clusters accompanied with significantly different means were then analysed further, especially those clusters with extreme high/low patient numbers.

It was found for the pollen season of *Ambrosia* that patient numbers are the highest in cluster 3 for each age category, most probably due to the highest and medium levels of *Ambrosia* and the remaining pollen, respectively (Table 2*a*). Cluster 4 involving a substantial part of summer provides the lowest patient numbers. It may be accounted for by moderate or low levels of both the two pollen types and the chemical pollutants (except for O_3) promoted by the highest wind speed (Table 2*a*). As regards the pollen-free season the highest patient numbers for each age category are associated with cluster 4. This can be explained by the relatively high temperature favourable for reproducing bacteria and viruses, as well as by strong winds that encourage the inflammation in the respiratory tracts by desiccating the air. Cluster 3 having anticyclonic character exhibits the lowest patient numbers for each age category, probably due to the very low temperatures in winter time that contribute to restrict respiratory infections (Table 2*b*).

3.2 Optimal time lags

Although there are examples for time lags even up to 8 days (Nascimento et al., 2006), the typical delays are up to 3 days in patient response to pollution exposure (e.g. Alves et al., 2010). It is likely that the explanatory variables express their effects in the formation of the respiratory problems within 3 days (Knight et al., 1991).

For instance, immediate allergic reactions of pollen can occur within 15-20 minutes, in certain cases 8-10 hours, while all immune reactions in cells can occur 48-72 hours following exposure (Petrányi, 2000). Our optimal time lag varies from zero to three days. With increasing age there is a tendency for more non-zero lags.

Table 2a. **Cluster-related mean values of the meteorological and pollutant parameters as well as patient numbers for the pollen season of *Ambrosia* (bold: maximum; *italic:* minimum)**

Cluster	1	2	3	4	5
Parameter	**Mean values**				
Total number of days	68	41	*26*	94	**137**
Frequency (%)	18.6	11.2	*7.1*	25.7	**37.4**
Temperature (°C)	23.3	16.9	20.5	**24.9**	*16.4*
Global solar flux ($W{\cdot}m^{-2}$)	211.1	155.3	176.4	**223.6**	*126.7*
Relative humidity (%)	66.1	72.2	68.6	*59.3*	**75.0**
Air pressure (hPa)	1002.7	**1009.4**	*1001.7*	1005.3	1005.9
Wind speed ($m{\cdot}s^{-1}$)	0.8	*0.5*	0.9	**1.1**	0.9
CO ($\mu g{\cdot}m^{-3}$)	468.9	**700.5**	*425.1*	444.5	463.8
PM_{10} ($\mu g{\cdot}m^{-3}$)	*36.3*	**52.8**	40.4	44.0	40.2
NO ($\mu g{\cdot}m^{-3}$)	10.8	**44.7**	14.1	*9.5*	15.1
NO_2 ($\mu g{\cdot}m^{-3}$)	32.9	**48.8**	33.2	34.0	*31.8*
O_3 ($\mu g{\cdot}m^{-3}$)	41.8	*26.2*	36.3	**58.4**	29.2
SO_2 ($\mu g{\cdot}m^{-3}$)	*4.0*	5.5	4.9	4.9	**6.0**
Ambrosia ($pollen{\cdot}m^{-3}{\cdot}day^{-1}$)	91.7	*43.3*	**593.2**	46.2	57.9
Total pollen excluding *Ambrosia*	**111.9**	16.8	48.7	49.9	*14.1*
Adults (15-64 years)	101.6	76.7	**114.3**	*74.5*	78.2
The elderly (≥65 years)	12.8	11.8	**13.3**	*10.5*	12.2
All age groups	114.5	88.7	**127.9**	*85.1*	90.6

Table 2b. **Cluster-related mean values of the meteorological and pollutant parameters as well as patient numbers for the pollen-free season (bold: maximum; *italic:* minimum)**

Cluster	1	2	3	4
Parameter	**Mean values**			
Total number of days	75	**137**	*44*	108
Frequency (%)	20.6	**37.6**	*12.1*	29.7
Temperature (°C)	**10.8**	3.2	*-2.7*	5.7
Global solar flux ($W \cdot m^{-2}$)	**62.1**	38.9	*36.2*	44.8
Relative humidity (%)	*82.1*	87.6	**93.1**	87.1
Air pressure (hPa)	1010.6	*1004.3*	**1020.3**	1008.5
Wind speed ($m \cdot s^{-1}$)	0.6	*0.5*	*0.5*	**1.4**
CO ($\mu g \cdot m^{-3}$)	788.2	**812.4**	729.7	*652.2*
PM_{10} ($\mu g \cdot m^{-3}$)	**79.2**	53.2	61.6	*52.6*
NO ($\mu g \cdot m^{-3}$)	35.9	**40.6**	*28.0*	31.9
NO_2 ($\mu g \cdot m^{-3}$)	**40.4**	38.0	*5.2*	36.0
O_3 ($\mu g \cdot m^{-3}$)	**19.6**	15.2	16.8	*11.3*
SO_2 ($\mu g \cdot m^{-3}$)	10.4	*6.6*	**15.2**	7.4
Adults (15-64 years)	61.4	59.2	*50.2*	**64.6**
The elderly (≥65 years)	11.7	11.6	*10.1*	**13.9**
All age groups	73.2	71.0	*60.3*	**78.5**

Table 3a. **Special transformation. Effect of the explanatory variables on respiratory diseases as resultant variables and the rank of importance of the explanatory variables on their factor loadings transformed to Factor 1 for determining the resultant variable; pollen season of *Ambrosia* (thresholds of significance: *italic:* $x_{0.05}$ = 0.056; bold: $x_{0.01}$ = 0.074)**

Explanatory variables	Adults (15-64 years)		The elderly (≥ 65 years)		All age groups	
	weight	**rank**	**weight**	**rank**	**weight**	**rank**
Patient number	**0.914**	–	**-0.983**	–	**0.918**	–
Temperature (°C)	**0.081**	10	0.021	13	0.067	10
Global solar flux ($W \cdot m^{-2}$)	**0.168**	7	*0.134*	4	**0.146**	8
Relative humidity (%)	-0.163	8	0.057	10	**-0.152**	7
Air pressure (hPa)	-0.042	13	-0.103	7	-0.018	13
Wind speed ($m \cdot s^{-1}$)	**-0.220**	5	**0.142**	2	**-0.228**	5
Total weight	0.675	–	0.457	–	0.611	–
CO ($\mu g \cdot m^{-3}$)	**-0.255**	3	0.095	8	**-0.251**	4
PM_{10} ($\mu g \cdot m^{-3}$)	**-0.314**	2	*0.119*	5	**-0.304**	2
NO ($\mu g \cdot m^{-3}$)	-0.058	11	-0.032	12	-0.041	12
NO_2 ($\mu g \cdot m^{-3}$)	0.047	12	*-0.134*	3	0.059	11
O_3 ($\mu g \cdot m^{-3}$)	**-0.230**	4	**0.272**	1	**-0.252**	3
SO_2 ($\mu g \cdot m^{-3}$)	*-0.115*	9	0.070	9	*-0.119*	9
Total weight	1.018	–	0.722	–	1.025	–
Ambrosia ($pollen \cdot m^{-3} \cdot day^{-1}$)	**0.553**	1	*-0.117*	6	**0.520**	1
Total pollen excluding *Ambrosia*	**0.199**	6	-0.041	11	**0.177**	6
Total weight	0.752	–	0.158	–	0.697	–

The global solar flux has the largest number of positive time shifts from meteorological variables (typically 2-3 days) for the elderly, while the relative humidity has the largest number of non-zero delays (2-3 days) for adults. However, the role of relative humidity in positive delays is substantially smaller than the role of the global solar flux. Within the chemical pollutants, positive lags (0-3 days) are mostly associated with CO and SO_2 for both age groups, and then with NO for adults and PM_{10} for the elderly, in agreement with other studies (e.g. Orazzo et al., 2009). No time shift is typical for the pollen season of *Ambrosia* in any allergy-sufferer group.

3.3. Factor analysis and special transformation

After performing a factor analysis for adults, the elderly and all age groups for the two seasons (altogether 3x2=6 factor analyses), 7 and 6 factors were retained for the pollen season of *Ambrosia* and pollen-free season, respectively. In order to calculate the rank of importance of the explanatory variables for determining the resultant variable, loadings of the retained factors were projected onto Factor 1 (with a special transformation) (Table 3a-b) (Jahn and Vahle, 1968).

As regards meteorological variables, only the wind speed and temperature display significant associations with hospital admissions for all three age groups in the pollen season of *Ambrosia*, which is confirmed by Freitas et al. (2010). The wind speed exhibits a dual character because strong winds facilitate decreasing hospital admissions by reducing the levels of the pollutants throughout the year, but they help to desiccate the air and hence encourage respiratory problems. The latter effect seems to be higher in the pollen-free season since the wind speed varies proportionally with the elderly patient numbers (Table 3b) in this period. In the pollen season of *Ambrosia*, the inverse relationship of wind speed with the number of hospital admissions suggests that the pollutant-diluting effect of the wind is predominant (Table 3a) here. In the cold pollen-free season, relatively high temperatures and strong wind speeds are associated with typical cyclonic air masses. This kind of weather can cause an increase in the number of hospital admissions of the elderly with weak immune systems, since their repeated exposition to air desiccated by winds may lead to an inflammation of the respiratory tracts (Table 3*b*) (Strausz, 2003).

Both for the pollen season of *Ambrosia* and pollen-free season, the total weight of the chemical pollutants is the highest for all three age groups with all variable types and is substantially higher than that of the meteorological variables. This latter finding may be due to the fact that anticyclonic weather situations, being the most frequent during the above seasons, favour enrichment of chemical pollutants, and so high pollutant levels have a greater effect on respiratory hospital admissions (Table 3*a-b*).

For the pollen season of *Ambrosia*, pollen variables display the second highest weight for adults and all age groups, mainly due to the very high *Ambrosia* pollen levels (Table 3a). For adults, the first three explanatory variables that influence patient numbers the most are, in decreasing order, *Ambrosia*, PM_{10} and CO, while for all age groups they are *Ambrosia*, PM_{10} and O_3, respectively. For the elderly, there are fewer significant associations between the explanatory variables and the number of respiratory problems; furthermore, pollen variables exhibit the smallest total weight in this case. Here, influencing variables ranked highest in decreasing order are O_3, wind speed and NO_2. *Ambrosia* pollen is still significantly correlated with the number of respiratory admissions, but it is ranked only 6 (Table 3*a*).

For the pollen-free season, the chemical variables are ranked highest (Table 3*b*). The sequence of the most important influencing variables in decreasing order for adults is NO_2, temperature and PM_{10}, for the elderly it is O_3, wind speed and NO_2, while for all age groups it is NO_2, temperature and O_3 (Table 3*b*).

Table 3b. **Special transformation. Effect of the explanatory variables on respiratory diseases as resultant variables and the rank of importance of the explanatory variables on their factor loadings transformed to Factor 1 for determining the resultant variable; pollen-free season (thresholds of significance: *italic:* $x_{0.05}$ = 0.105; bold: $x_{0.01}$ = 0.138)**

Explanatory variables	Adults (15-64 years)		The elderly (≥ 65 years)		All age groups	
	weight	rank	weight	rank	weight	rank
Patient number	**0.993**	–	**0.943**	–	**0.992**	–
Temperature (°C)	**0.168**	2	**0.165**	4	**0.188**	2
Global solar flux ($W \cdot m^{-2}$)	0.048	8	-0.004	11	0.047	9
Relative humidity (%)	0.009	10	-0.083	9	-0.020	11
Air pressure (hPa)	*0.110*	7	*0.124*	7	0.055	7
Wind speed ($m \cdot s^{-1}$)	0.035	9	**0.273**	2	0.022	10
Total weight	0.370	–	0.649	–	0.331	–
CO ($\mu g \cdot m^{-3}$)	-0.009	11	*-0.126*	6	-0.049	8
PM_{10} ($\mu g \cdot m^{-3}$)	**0.147**	3	0.092	8	**0.138**	4
NO ($\mu g \cdot m^{-3}$)	*0.120*	6	**0.148**	5	0.104	5
NO_2 ($\mu g \cdot m^{-3}$)	**0.175**	1	**0.186**	3	**0.197**	1
O_3 ($\mu g \cdot m^{-3}$)	*-0.121*	5	**-0.351**	1	**-0.158**	3
SO_2 ($\mu g \cdot m^{-3}$)	*-0.124*	4	-0.073	10	-0.091	6
Total weight	0.696	–	0.977	–	0.737	–

3.4. Time-varying multivariate linear regression

The ratio of variance of patient numbers accounted for by explanatory variables to variance of patient numbers is considerably larger for adults than for elderly people (Table 4). Also, the annual cycle of adult patient numbers is substantially larger (Table 4), suggesting that hospital visits of elderly people depend on pollutants and meteorological conditions to a lesser degree. This may be due to the following social factor. Some of the elderly habits tend to underestimate chronic diseases and consider them as a natural attendant of age. Hence, the elderly often do not turn to physician and seek medical treatment in time (Johnson, 2005).

Here, three questions should be raised. First, what is the order of explanatory variables in view of the strength of the influence on patient number? Second, what is the relative contribution of these variables to the variance explained by all of the variables? And third, how do the patient numbers change with a unit change in the explanatory variables? Regarding the first question, the selection of importance of explanatory variables in the formation of patient numbers was performed by the well-known stepwise regression method (Draper and Smith, 1981).

Table 4. **Relative variance (%) of patient numbers accounted for by explanatory variables including and omitting (in parentheses) the annual cycle of patient numbers**

Patients	**Pollen season of *Ambrosia***	**Pollen-free season**
Adults	26.8 (50.0)	17.8 (21.2)
The elderly	19.9 (23.9)	9.9 (13.9)
All age groups	26.5 (48.1)	13.8 (17.7)

The answer to the second question is more difficult due to the multicollinearity among the explanatory variables. Namely, the sum of variances explained by individual variables is larger than the variance explained by all of the variables. Therefore, neither univariate regressions with individual explanatory variables nor the multivariate regression using all of the variables is appropriate to quantify the individual explained variances. However, elementary consideration shows that the variance percentage explained by the ith variable lies between V-V_i and v_i, where V is the total explained variance, and v_i and V_i are the variances explained by the ith variable and all of the variables excluding the ith variable, respectively. Therefore, only the ranges of variances accounted for by explanatory variables are shown in Table 5. The most important explanatory variables that influence patient numbers in the pollen season of *Ambrosia* are O_3 and temperature for each age category. The order of the remaining variables is the same for adults and all of ages: global solar flux, NO, wind speed, PM_{10}, air pressure, and *Ambrosia*. Almost the same variables are considered to be most significant for the elderly group, but with a slightly different order. Surprisingly, *Ambrosia* is only the sixth-eighth most important explanatory variable influencing patient numbers. For the pollen-free season, O_3 is the second main factor for adults and all age groups, but is just fifth for the elderly. The order varies with the different age groups. For instance, the most significant variable is PM_{10}, NO and global solar flux for adults, the elderly and all age categories, respectively.

In addition, the order of importance of the explaining variables identified by the stepwise regression method is not the same as the order of level of the statistical significance (Table 5). The significance depends not only on the strength of the relationship, but also on data length and autocorrelations of the different variables. Significance levels were determined by a Monte-Carlo simulation experiment. Approximating the autocorrelations of an explaining variable by a first order autoregressive model fitted to observed values of this variable, a time series independent of patient numbers was generated according to the time-varying empirical probability distribution function of the underlying explaining variable. The observed values were then substituted by these simulated data and a time-varying multivariate linear regression was performed. Then the mean squared error for patient numbers obtained from this regression was calculated. These steps were repeated 1,000 times, and appropriate quantiles of the empirical probability distribution function of these 1,000 simulated mean squared errors yielded the critical value for checking the null-hypothesis of being this explaining variable uncorrelated with patient numbers. The procedure was applied to each explaining variable separately.

The variation of the patient numbers with unit changes in the explanatory variables exhibit annual cycles as the regression coefficients depend on dates within the year. There is evidence that confirms different effects of the explanatory variables in different periods of the year.

Table 5. **Ratio (%) of variances accounted for by explanatory variables to variance accounted for by all of the explanatory variables. Variables are indicated in an order of importance obtained via a stepwise regression. Only variables with a joint contribution just exceeding 90% of the total explained variance are shown. (in *X:* significant for $p < 0.1$, in X: significant for $p < 0.05$, in *X*: significant for $p < 0.01$)**

Pollen season of *Ambrosia*			**Pollen-free season**		
Adults	**The elderly**	**All ages**	**Adults**	**The elderly**	**All ages**
O_3: 15.8-31.7	**O_3: 20.1-36.7**	***O_3: 16.8-46.6***	**PM_{10}: 14.6-18.5**	*NO: 10.0-17.1*	**GSF: 8.0-12.3**
T: 1.5-10.9	T: 1.0- 5.0	*T: 1.5-10.4*	*O_3: 7.9- 9.0*	CO: 8.0-18.1	*O_3: 10.9-12.3*
GSF: 2.3- 6.4	NO_2: 5.0-10.0	GSF: 2.2- 6.0	*RH: 3.4-19.1*	SO_2: 8.0-14.1	*RH: 4.3-20.3*
NO: 3.8- 6.4	WS: 5.0- 6.0	NO: 3.7- 6.3	T: 7.9- 8.4	RH: 3.0-16.1	*CO: 4.3-11.6*
WS: 6.8-10.2	PM_{10}: 3.5-10.1	*WS: 7.1-10.1*	P: 0.2- 9.6	*O_3: 13.1-16.1*	*NO_2: 12.3-13.8*
PM_{10}: 6.0- 7.9	A: 2.5- 4.0	PM_{10}: 6.0- 8.2	*GSF: 8.4-10.7*	WS: 5.0- 9.0	P: 0.1- 9.4
P: 0.1- 5.3	GSF: 4.5- 5.0	P: 0.1- 5.0	SO_2: 5.6- 6.2	NO_2: 8.0-13.1	*NO: 8.0-16.7*
A: 5.7- 6.4	P: 0.0- 3.5	A: 4.9- 6.3	*NO: 7.3-12.9*	PM_{10}: 5.0- 5.1	T: 5.1- 9.4
RH: 1.1- 4.2	CO: 2.0- 4.0	RH: 0.7- 3.7			

T = Temperature (°C); *GSF* = Global Solar Flux (W·m^{-2}); *RH* = Relative humidity (%); *P* = Air pressure (hPa); WS = Wind speed (m·s^{-1}); *CO* = Carbon-monoxide (μg·m^{-3}); PM_{10} = Particulate matter smaller size than 10 μm (μg·m^{-3}); *NO* = Nitrogen-monoxide (μg·m^{-3}); NO_2 = Nitrogen-dioxide (μg·m^{-3}); O_3 = Ozone (μg·m^{-3}); SO_2 = = Sulphur-dioxide (μg·m^{-3}); *A* = *Ambrosia* (pollen·m^{-3}·day^{-1}).

Table 6. **Minima and maxima of regression coefficients during the year**

Variable	Pollen season of Ambrosia						Pollen-free season					
	Minimum			Maximum			Minimum			Maximum		
	Adults	The elderly	All	Adult	The elderly	All	Adults	The elderly	All	Adults	The elderly	All
T	-3.61	-0.27	-3.96	5.28	0.01	5.83	-0.52	-0.55	-0.59	2.26	0.08	1.58
GSF	-0.03	-0.02	-0.03	0.38	0.01	0.40	-0.20	-0.02	-0.13	0.13	0.00	0.13
RH	-0.54	-0.06	-0.56	0.17	0.01	0.18	-0.77	-0.21	-0.82	0.29	0.05	0.35
P	-0.04	-0.01	-0.05	0.01	0.01	0.07	-0.03	0.00	-0.03	0.06	0.02	0.08
WS	-35.1	-2.61	-37.7	0.26	0.33	0.17	-4.11	0.00	-3.50	4.49	3.72	6.06
CO	-0.04	-0.01	-0.04	0.00	0.01	0.00	-0.01	-0.01	-0.02	0.01	0.01	0.00
PM_{10}	-0.50	-0.03	-0.55	0.14	0.03	0.18	0.00	-0.04	-0.05	0.24	0.04	0.13
NO	-1.00	-0.11	-1.07	0.25	0.01	0.23	-0.15	-0.03	-0.18	0.27	0.09	0.38
NO_2	-0.26	-0.07	-0.31	0.13	0.09	0.16	-0.42	-0.14	-0.68	0.21	0.08	0.47
O_3	-1.20	-0.16	-1.37	0.60	0.03	0.10	-0.47	-0.20	-0.70	0.24	0.02	0.00
SO_2	-1.54	-0.20	-1.75	0.63	0.18	0.99	-1.24	-0.09	-0.54	0.00	0.21	0.47
A	-0.08	-0.03	-0.11	2.78	0.24	3.05						
TP	-0.02	-0.13	0.00	0.34	0.02	0.59						

T = Temperature (°C); *GSF* = Global Solar Flux ($W \cdot m^{-2}$); *RH* = Relative humidity (%); *P* = Air pressure (hPa); *WS* = Wind speed ($m \cdot s^{-1}$); *CO* = Carbon-monoxide ($\mu g \cdot m^{-3}$); PM_{10} = Particulate matter smaller size than 10 μm ($\mu g \cdot m^{-3}$); *NO* = Nitrogen-monoxide ($\mu g \cdot m^{-3}$); NO_2 = Nitrogen-dioxide ($\mu g \cdot m^{-3}$); O_3 = Ozone ($\mu g \cdot m^{-3}$); SO_2 = Sulphur-dioxide ($\mu g \cdot m^{-3}$); *A* = *Ambrosia* ($pollen \cdot m^{-3} \cdot day^{-1}$); *TP* = Total pollen ($pollen \cdot m^{-3} \cdot day^{-1}$)

For instance, the wind speed is inversely proportional to the patient numbers in the pollen season of *Ambrosia* due to the diluting effect of wind. In the pollen-free season, however, the wind speed is mainly in positive association with respiratory problems, especially for the elderly with weak immune systems as repeated exposition to strong winds may aid the inflammation of the respiratory tracts (Strausz, 2003). Furthermore, the optimal conditions for bacteria and viruses affecting respiratory problems are different. While, for example, Mycoplasma bacteria generating pneumonia and other respiratory inflammations favour low relative humidity (pollen season of *Ambrosia*), adenoviruses provoking upper respiratory infections and conjunctivitis are more infectious at a higher relative humidity (pollen-free season) (Strausz, 2003). The minima and maxima of regression coefficients (Table 6) tell us the boundaries of the mean patient number change with a unit change in different explanatory variables.

ACKNOWLEDGEMENTS

The authors would like to thank Gábor Motika (Environmental Conservancy Inspectorate, Szeged, Hungary) for supplying meteorological data for Szeged and Miklós Juhász (University of Szeged) for providing pollen data for Szeged. This study was supported by the TRANSAIRCULTUR project (No. HURO/1001/139/1.3.4) Hungary-Romania Cross-Border Co-operation Programme 2007-2013, under the auspices of the European Union and the European Regional Development Fund.

REFERENCES

[1] Ågren C., 2010. *Particles killing half a million.* Acid News 2, 1-5. AirClim. Air Pollution & Climate Secretariat, Sweden. http://www.airclim.org/acidnews/2010/AN2-10.php#1

[2] Alves C.A., Scotto M.G., Freitas M.C., 2010. *Air Pollution and Emergency Admissions For Cardiorespiratory Diseases in Lisbon (Portugal).* Quimica Nova 33(2), 337-344.

[3] Cohen A.J., Anderson H.R., Ostro B., Pandey K.D., Krzyzanowski M., Künzli N., Gutschmidt K., Pope A., Romieu I., Samet J.M., Smith K., 2005. *The global burden of disease due to outdoor air pollution.* Journal of Toxicology and Environmental Health A 68(13-14), 1301-1307.

[4] D'Amato G., Liccardi G., D'Amato M., Holgate S., 2005. *Environmental risk factors and allergic bronchial asthma.* Clinical and Experimental Allergy 35(9), 1113-1124.

[5] Draper N., Smith H., 1981. *Applied Regression Analysis,* 2d Edition, John Wiley & Sons, New York.

[6] Freitas M.C., Pacheco A.M.G., Verburg T.G., Wolterbeek H.T., 2010. *Effect of particulate matter, atmospheric gases, temperature, and humidity on respiratory and circulatory diseases' trends in Lisbon,* Portugal. Environmental Monitoring and Assessment 162(1-4), 113-121.

[7] Galán C., Cariňanos P., García-Mozo H., Alcázar P., Domínguez-Vilches E., 2001. *Model for forecasting Olea europaea L. airborne pollen in South-West Andalusia, Spain.* International Journal of Biometeorology 45(2), 59-63.

[8] Jahn W., Vahle H., 1968. *Die Faktoranalyse und ihre Anwendung.* Verlag die Wirtschaft, Berlin, 231 pp. (in German)

[9] Johnson M.L., (Ed), 2005. *The Cambridge Handbook of Age and Ageing,* Cambridge University Press.

[10] Jolliffe I.T., 1993. *Principal component analysis: A beginner's guide – II. Pitfalls, myths and extensions.* Weather 48, 246-253.

[11] Just J., Nikasinovic L., Laoudi Y., Grimfeld A., 2007. *Air pollution and asthma in children.* Revue Francaise d'Allergologie et d'Immunologie Clinique 47(3), 207-213.

[12] Knight A., Drouin M.A., Yang W.H., Alexander M., Delcarpio J., Arnott W.S., 1991. *Clinical-evaluation of the efficacy and safety of noberastine, a new H1 antagonist, in seasonal allergic rhinitis - a placebo-controlled, dose-response study.* Journal of Allergy and Clinical Immunology 88(6), 926-934.

[13] Liu, P.W.G., 2009. *Simulation of the daily average PM10 concentrations at Ta-Liao with Box-Jenkins time series models and multivariate analysis.* Atmospheric Environment 43, 2104-2113.

[14] Mahalanobis P.C., 1936. *On the generalised distance in statistics*. Proceedings of the National Institute of Sciences of India 12, 49-55.

[15] Makra L., Sánta T., Matyasovszky I., Damialis A., Karatzas K., Bergmann K.C., Vokou D., 2010. *Airborne pollen in three European cities: Detection of atmospheric circulation pathways by applying three-dimensional clustering of backward trajectories.* Journal of Geophysical Research-Atmospheres 115, D24220, doi:10.1029/2010JD014743

[16] Nascimento L.F.C., Pereira L.A.A., Braga A.L.F., Modolo M.C.C., Carvalho J.A., 2006. *Effects of air pollution on children's health in a city in Southeastern Brazil.* Revista de Saude Publica 40(1), 77-82.

[17] Orazzo, F., Nespoli, L., Ito, K., Tassinari, D., Giardina, D., Funis, M., Cecchi, A., Trapani, C., Forgeschi, G., Vignini, M., Nosetti, L., Pigna, S., Zanobetti, A., 2009. *Air Pollution, Aeroallergens, and Emergency Room Visits for Acute Respiratory Diseases and Gastroenteric Disorders among Young Children in Six Italian Cities.* Environmental Health Perspectives 117(11), 1780-1785.

[18] Petrányi Gy., (Ed.), 2000. *Klinikai Immunológia* (Clinical Immunology). Medicina Könyvkiadó, Budapest, ISBN: 9789632425610 (in Hungarian)

[19] Strausz, J., (Ed.), 2003. *Asthma bronchiale.* Mediszter GlaxoSmithKline, Budapest.

[20] Tukey J.W., 1985. *The problem of multiple comparisons (1953)* (unpublished manuscript), in: The Collected Works of John W. Tukey. Vol. II: Time Series, pp. 1965-1984. Wadsworth Advanced Books & Software, Monterey, CA.