Analysis of Fraxinus pollen seasons and forecast models based on meteorological factors

Agnieszka Kubik-Komar1,A,C,F, Krystyna Piotrowska-Weryszko2,A,B,D,F, Elżbieta Weryszko-Chmielewska2,B,D,F, Bogusław Michał Kaszewski3,B,E,F

1 Department of Applied Mathematics and Computer Science, University of Life Sciences, Lublin, Poland
2 Department of Botany, University of Life Sciences, Lublin, Poland
3 Department of Meteorology and Climatology, Maria Skłodowska-Curie University, Lublin, Poland

Abstract

Introduction and objective. The timings of Fraxinus and Betula flowering and pollen release overlap, which may cause increased allergic reactions in sensitive people. The aim of the present study was to characterize Fraxinus pollen seasons in Lublin (central-eastern Poland) and to identify meteorological factors that most determine the occurrence of airborne pollen of this taxon, as well as obtain forecast models for the basic characteristics of the pollen season.

Materials and method. The study was conducted in Lublin during the period 2001–2016, employing the volumetric method. The seasons were compared by PCA (Principal Component Analysis). To determine relationships between meteorological conditions and the pattern of pollen seasons, regression analysis was used. Data for the period 2001–2015 were used to create forecast models by applying regression analysis, while the 2016 data served to verify these models.

Results. Season end date and seasonal peak date were characterized by the lowest variation. The biggest differences were found for peak value and total annual pollen sum. The average dates of occurrence of ash pollen grains in the air of Lublin were between 13 April – 3 May, whereas, on average, the pollen peak date occurred on 23 April. The factor loading values for the PC1 variable indicate that it is most strongly correlated with peak value and total pollen sum, while the PC2 variable correlated with the pollen season start date and season duration (a negative correlation). Regression models were developed for the following pollen season characteristics: season start, end and duration, seasonal peak date, and total annual pollen sum.

Conclusions. The fit of the forecast models was at the level of 62–94%. Analysis of the data showed that weather conditions mainly in February were important factors controlling the Fraxinus pollen season.

Key words

allergy, regression analysis, aerobiology, airborne pollen, Fraxinus

INTRODUCTION

Trees of the genus Fraxinus (f. Oleaceae) comprising 65 species grow in temperate areas of the Northern Hemisphere [1]. In Poland, the only native species is Fraxinus excelsior L. This species occurs in central and southern Europe, in the Caucasus and Asia Minor [1, 2, 3].

F. excelsior is frequently found in moist forests and river valleys. It is also planted in parks, and along avenues and roads. For urban plantings, other species of this genus are also used: F. americana and F. pennsylvanica originating from North America, as well as F. ornus and F. angustifolia from southern Europe [1]. In Poland, the flowering of F. excelsior occurs in April and May. The average number of pollen grains produced by one flower of this species is 25,050, while for one inflorescence it is 1,605,700 [4].

In the Oleaceae family, Fraxinus is one of the sources of pollen allergens. Olea europaea belongs to species of major importance in the etiology of allergic diseases in the Mediterranean region [5]. Because one of the main Fraxinus pollen allergens, Fra e 1 with a molecular mass of 7 kDa, exhibits homology with an Olea pollen allergen, Ole e 1 (16 kDa), in patients with hypersensitivity to ash pollen allergens there is a risk of cross-reactions [6, 7, 8].

The possibility of occurrence of cross-reactions in allergy sufferers sensitive to Fraxinus pollen also applies to other species belonging to Oleaceae and planted as ornamental shrubs: Ligustrum, Forsythia, Syringa, Jasminum [7, 8, 9].

In central Europe the main pollen season of Fraxinus largely overlaps with the Betula pollen shed period. Betula pollen causes allergy in a substantial part of pollen-sensitive people in Poland (14.9%) [10]. Research also confirms that there is a risk of cross-reactions between Betula and Fraxinus pollen allergens [7]. The above-mentioned author also reports that patients allergic to Alnus and Corylus pollen are also frequently sensitized to Fraxinus pollen. The possible incidence of allergies to Fraxinus pollen is aggravated by the presence of allergens in orbicules on the exine surface and on the inner surface of anthers of the above-mentioned taxon during pollen release [11].

Research on Fraxinus in Poland can be found in [12] and [13] where the concentration comparison of this taxon in different Polish regions is presented, as well as in [14] where the hourly distribution of pollen count and the correlation between pollen concentration of Fraxinus and some...
environmental factors were studied. However, the risk of allergy to Fraxinus is still scarcely reported. Some authors have determined the threshold values for Fraxinus pollen concentrations that cause allergy symptoms in the average patient. Horak et al. [15] report that for the Czech Republic this concentration is 167 P/m³. In Switzerland, on the other hand, the threshold value was determined to be 100 P/m³, while in Austria 90 P/m³, but these values were determined without investigating allergic symptoms in patients. The reported values are within the limits of moderate (16–90) and high (91–1500) tree pollen concentrations, as determined by Frenz [16]. In this study, it was assumed that high (> 90 P/m³) daily pollen concentrations [16] may be a risk of allergy in susceptible individuals.

Multivariate methods are widely used in aerobiological studies, the most popular of which are PCA, cluster analysis and regression. The latter is commonly used in forecast modeling [17, 18, 19]. In this study, stepwise multiple regression was applied [5]. The application of PCA as well as cluster analysis in aerobiological studies can be found in articles [20, 21].

OBJECTIVES

Despite the articles mentioned above, information on the Fraxinus pollen season in Poland is still limited and insufficient. Therefore, the aim of this study was to compare the pattern of Fraxinus pollen seasons in Lublin during the period 2001–2016, to determine the effects of meteorological factors on their dynamics, and to develop regression models to forecast the major parameters of the Fraxinus pollen season. In assessment of the risk of pollen allergy, it is important to determine the number of days of high pollen concentration during a season, which is more than 90 P/m³ for trees. The number of days on which there was such a risk was determined for individual years.

MATERIAL AND METHODS

This study on monitoring airborne Fraxinus pollen grains was conducted in Lublin (central-eastern Poland) during the period 2001–2016. Average daily pollen concentrations were measured with a Hirst-type pollen trap (Lanzoni VPPS 2000). The pollen trap was located on the flat roof of a building of the University of Life Sciences at an altitude of 197 m, and at a height of 18 m above ground level. Standard methods applied in aerobiological research were employed for analysis [22]. Results were expressed as the number of pollen grains per 1 m² of air per 24 hours (P/m²). The following parameters of the pollen season were analyzed: start, end, duration, maximum daily pollen concentration (peak value), maximum daily concentration date (peak date), and total annual pollen sum of average daily pollen concentration. The 95% method was used to calculate the start and end of the pollen seasons [23].

Principal component analysis (PCA) was used to compare the pollen seasons, at the same time taking into account the individual season characteristics. 15 pollen seasons and 6 season characteristics were treated as a data matrix. The aim of this analysis was to find two new parameters that would allow the pollen season characteristics in individual years to be presented graphically on one plot. Factor loading values were determined after a normalized varimax rotation so as to ensure that the new variables, PC1 and PC2, would provide the highest possible percentage of the explained variance. The factor scores shown in the plot, whose axes were the determined variables, allowed quick identification of the years with similar values of the season characteristics.

Ash pollen seasons were also analyzed in relation to weather conditions. The meteorological data came from the Meteorological Observatory of the Meteorology and Climatology Department, the Maria Curie-Skłodowska University in Lublin, which is located at a distance of about 1.5 km from the pollen sampling site. The following meteorological data were used for the analysis: mean, minimum and maximum air temperature, relative air humidity, rainfall, cloud cover, and wind speed. To determine relationships between meteorological conditions and season characteristics, Spearman’s correlation coefficient was applied, while at a further stage, regression analysis was employed. The 2001–2015 data were used as a learning sample to find regression models for the individual season characteristics. 10-day data for the meteorological factors were used as independent variables; the data matrix therefore consisted of 15 seasons and 72 variables (4 months x 3 decades x 6 meteorological factors). The wind speed was not taken into consideration due to some missing values. Similar to the study by Stach et al. [17], to prevent multicollinearity, it was decided not to enter highly correlated independent variables into the multiple regression models. For example, from maximum, mean and minimum temperatures of the same 10-day period, only one of these parameters was included in the model (if any). In seeking optimal functions that would reflect between-year variations in the characteristics of the seasons, a multiple and polynomial regression was conducted, entering stepwise the individual variables into the model and thus building step by step the best regression model [24, 25]. The best-fitting regression model was chosen based on the value of the adjusted coefficient of determination (AdjR²). The predictive power was verified by calculating the model’s values for the 2016 data and comparing them with the actual values.

RESULTS

On average, the Fraxinus pollen season started on 13 April (Tab. 1). The difference in the extreme season start dates was 21 days over the study period. The earliest season onset was recorded in 2002 (2 April) and the latest in 2006 and 2013 (23 April). The coefficient of variation of the characteristic in question was 6.4%. It is worth stressing that in the last 3 years (2014–2016) the season began earlier than the 16-year average would indicate. Season end date was characterized by much lower variations (V=2.7%). On average, the season end was recorded on 3 May, with the earliest date on 26 April (2014) and the latest on 7 May (2006) (Fig. 1). The ash pollen season lasted, on average, 3 weeks, and the difference between the shortest and longest season was 16 days. The longest pollen seasons occurred in the years 2001, 2002 and 2005 (on average, 29 days), but in most years during the period 2003–2013 their average duration was much shorter (18 days). In the years 2014–2016, however, there was an extension in season duration (average season duration – 24 days). Seasonal peaks and total annual pollen sums were characterized by very high
Analysis of Fraxinus pollen seasons

The lowest peak value, which was recorded in 2016, was 6.6 times lower than the highest peak value observed in 2006 (Fig. 2). This plot shows that there is no linear trend in seasonal maximum. Among the years studied, 2006 was characterized by a short season, the latest start and end dates, as well as by the highest peak value and total annual pollen count (Tab. 1). High Fraxinus pollen concentrations exceeding the threshold value were recorded for a maximum of 12 days, whereas the average for the 16-year study period was half as small – 6 days. It was only in 2016 that no day with a pollen concentration exceeding the value of 90 P/m$^3$ was recorded (Tab. 1).

Statistical analysis revealed a strong negative correlation between start date and season duration, which means that a season that began earlier usually lasted longer, while a season that began late usually lasted for a shorter time. Peak value and annual total were found to be positively correlated. Correlation analysis also showed that a pollen season that started earlier also ended earlier, and that a season with an earlier end date was characterized by an earlier peak date (a positive correlation) (Tab. 2).

PCA allows reduction in the size of data sets to a much smaller number of independent principal components, explaining the biggest impact on total variability of the original data set, without much loss of information [26]. The determined factor loading values provided almost 82% of the explained variance of the entire structure by the 2 new variables (Tab. 3). The factor loading values for the PC1 variable indicate that it is most strongly correlated with peak value and annual total; therefore, this factor can be termed as the season's intensity. This variable is also quite largely correlates negatively with the peak value and the end of season. The PC2 variable is most strongly correlated to pollen season start date and season duration (negative correlation). The factor scores shown in the plot, whose axes were the PC1 and PC2, allowed quick identification of the years with similar values of the season characteristics (Fig. 3).

### Table 1. Statistics of the parameters of the Fraxinus pollen season in Lublin in 2001–2016

<table>
<thead>
<tr>
<th>Statistics</th>
<th>Pollen season Duration</th>
<th>Peak</th>
<th>Annual total</th>
<th>Day number with count above threshold</th>
</tr>
</thead>
<tbody>
<tr>
<td>SD</td>
<td>6.6 3.4 5.3 130.2 5.7 757.4</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>V (%)</td>
<td>6.4 2.7 25.8 45.6 5.1 42.3 56.8</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Table 2. List of significant Spearman’s correlations between the parameters of Fraxinus pollen season in Lublin (2001–2016)

<table>
<thead>
<tr>
<th>Parameters of pollen season</th>
<th>Spearman coefficient</th>
</tr>
</thead>
<tbody>
<tr>
<td>Start &amp; Duration</td>
<td>-0.9332**</td>
</tr>
<tr>
<td>Start &amp; End</td>
<td>0.4892*</td>
</tr>
<tr>
<td>End &amp; Peak Date</td>
<td>0.5801*</td>
</tr>
<tr>
<td>Peak Value &amp; Annual Total</td>
<td>0.9000**</td>
</tr>
</tbody>
</table>

Level of significance * 0.05, ** 0.01

### Table 3. Factor loadings after VARIMAX Rotation

<table>
<thead>
<tr>
<th></th>
<th>PC1</th>
<th>PC2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Start</td>
<td>0.004</td>
<td>0.986*</td>
</tr>
<tr>
<td>End</td>
<td>-0.548</td>
<td>0.617</td>
</tr>
<tr>
<td>Duration</td>
<td>-0.352</td>
<td>-0.838*</td>
</tr>
<tr>
<td>Peak value</td>
<td>0.919*</td>
<td>0.113</td>
</tr>
<tr>
<td>Peak date</td>
<td>-0.615</td>
<td>0.583</td>
</tr>
<tr>
<td>Annual total</td>
<td>0.918*</td>
<td>0.085</td>
</tr>
<tr>
<td>Explained variance</td>
<td>44.52%</td>
<td>37.24%</td>
</tr>
</tbody>
</table>

* factor loading value >0.8

---

Figure 1. Start and end dates of Fraxinus pollen season in Lublin, 2001–2016

Figure 2. Maximum daily concentration of Fraxinus airborne pollen in Lublin, 2001–2016.

Figure 3. Scatterplot of factor scores in PC1-PC2 coordinate system.
The location of the point corresponding to the 2006 season in the plot indicates that this was a specific season characterized by a late onset, a short duration, and at the same time, very abundant pollen production. Other seasons that started late and lasted relatively short, but with varying airborne pollen intensity, were the seasons in 2003, 2012 and 2013. The graph plot shows that the seasons in 2002, 2005, 2010 and 2014 began earliest and were characterized by a long duration. Moreover, it can be read from the plot that, apart from the year 2006, a high peak value and a high annual total were observed in the seasons in 2008, 2010, 2011, 2014 and 2015. The seasons in 2001, 2005, 2007, 2009 and 2012, characterized by a small amount of airborne pollen and a low peak value, are on the opposite side of the PCI axis. Additionally, the point for 2001 is the most to the left on the graph, also because, apart from the small concentration of pollen, it was characterized by the late peak and the end of the season. PCA demonstrates that the *Fraxinus* pollen seasons were most similar in the years 2008, 2010, 2014 and 2015, as well as in 2007 and 2009.

Based on the 2001–2015 data, regression models were developed for pollen season start, end and duration, as well as for peak date and annual total. The fit of the models was at the level of 62–94% (Tab. 4). Verification of the models was made based on the 2016 data and it so happened that the predicted and actual values were similar (e.g. start, end, duration, peak date – a one-day difference).

![Figure 4](image4.png)

**Figure 4.** The start date of the *Fraxinus* pollen season; actual vs. regression model calculated values and predicted value in 2016

Rainfall in the second 10 days of January, humidity in the first 10 days of February, and mean temperature in the second 10 days of February, largely determined *Fraxinus* pollen season duration. The above-mentioned variables explain 79% of the variation in this characteristic (Adj$R^2=0.79$). The difference between the values calculated from the regression equation and the actual values, on average, was 1.7 day. In 2016, the season end was recorded one day earlier than the date predicted on the basis of the regression model (Fig. 5).

![Figure 5](image5.png)

**Figure 5.** The end date of the *Fraxinus* pollen season; actual vs. regression model calculated values and predicted value in 2016

The best fit was obtained for pollen season start date because the regression equation explained the variation in season start in 94%. The forecast model for pollen season start was based on 2 variables: mean temperature in the second 10 days of February, and air humidity in the first 10 days of February. The difference between the values calculated from the regression equation and the actual values was one day, on average, whereas the greatest deviation between the actual and calculated values was 3 days (in 2011) (Fig.4). The predicted date in 2016 was a day earlier than the actual date.

The regression model describing season end included the following factors: mean temperature in the second 10 days of February, maximum temperature in the third 10 days of February, as well as cloud cover in the first 10 days of March and in the second 10 days of April. These variables explain 62% of the variation in season end dates (Adj$R^2=0.62$). The difference between the values calculated from the regression equation and the actual values, on average, was 1.3 day.

### Table 4. Regression models for *Fraxinus* pollen season parameters, based on meteorological factors in Lublin

<table>
<thead>
<tr>
<th>Season parameters</th>
<th>Regression equation</th>
<th>$R^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Start</td>
<td>$428.62-8.77H_{2}+0.46T_{mean_2}+0.13(T_{mean_3})^{2}$</td>
<td>0.94</td>
</tr>
<tr>
<td>End</td>
<td>$98.52-0.54T_{mean_3}+1.76C_{3}+1.82T_{mean_1}-0.48T_{mean_2}$</td>
<td>0.62</td>
</tr>
<tr>
<td>Duration</td>
<td>$40.42-0.61R_{3}^2-0.62H_{2}+1.38T_{mean_2}$</td>
<td>0.79</td>
</tr>
<tr>
<td>Peak date</td>
<td>$55.10+0.86T_{mean_3}+2.12T_{max_2}-0.15(T_{mean_1})^{2}-4.82R_{3}^{1}$</td>
<td>0.85</td>
</tr>
<tr>
<td>Annual total</td>
<td>$-1255.51+13.58(T_{mean_3})^{2}+3.96(T_{max_2})^{2}+35.24(T_{mean_1})^{2}$</td>
<td>0.90</td>
</tr>
</tbody>
</table>

$C_{3} =$ cloud cover in the third 10 days of February,
$H_{2} =$ humidity in the first 10 days of February,
$R_{3} =$ rainfall in the second 10 days of January,
$T_{mean_1} =$ mean temperature in the third 10 days of January,
$T_{max_2} =$ maximum temperature in the third 10 days of February,
The timing of *Fraxinus* pollen seasons in Lublin in particular years significantly varied. It should be stressed that the difference between the latest season start date (23 April) and the earliest end date (26 April) was only 3 days (Fig. 1). Air temperature proved to be the factor that most affected the timing of pollen seasons. In the years 2006 and 2013, with the latest onset of the *Fraxinus* pollen season, the mean monthly temperature in January, February and March was below 0 °C, whereas in 2002, in which the season onset was recorded earliest, the year was characterized by exceptionally warm February and March. A comparison of the data regarding the start dates for the *Fraxinus* and *Betula* pollen seasons in Lublin during 2001–2010, reveals that the start of the pollen seasons of these taxa occurred with similar timing, which may be of essential importance for pollen allergic people. A previous study demonstrates that cross-reactions to pollen allergens of these tree genera may occur [27, 28, 29].

**DISCUSSION**

*Fraxinus* pollen may cause spring pollinosis in areas where it occurs in substantial amounts. High concentrations of this taxon’s pollen are recorded in central [6, 27, 28, 29] and southern Europe [30, 31, 32]. The presented study shows that in Poland high *Fraxinus* pollen concentrations may pose a great threat to allergy sufferers. The results of this study conducted in Lublin reveal that maximum daily concentrations of *Fraxinus* pollen in central-eastern Poland reach even 545 P/m³, while the average maximum value for the 16-year study period is 286 P/m³. Analysis of the pollen seasons of *Fraxinus* in 8 Polish cities during 2001–2005 shows that the highest maximum pollen concentration of this taxon was in Lublin [33]. In the subsequent years of research (2007–2008), there were also higher concentrations of ash pollen in Lublin than in other Polish cities. Additionally, the peak values of this taxon were 1.4–4.4 times higher in Lublin than in other cities [34, 35].

In Poland, the maximum *Fraxinus* pollen concentrations measured at 8 pollen sampling sites ranged between 28–412 P/m³ during the period 2001–2005, and in most cases (64%) they reached high values [33]. In 2007, at 7 other monitoring sites, maximum values of 44–193 P/m³ were recorded, while in 2008 it was 125–454 P/m³, but in both these years the highest pollen concentrations were found in Lublin. In Poland, the average values for maximum *Fraxinus* pollen concentrations recorded at the above-mentioned pollen monitoring sites were 87 P/m³ in 2007 and 242 P/m³ in 2008 [34, 35].

In Cordoba and Plasencia, Spain, 59% and 68% respectively of patients with a diagnosis of pollinosis were sensitized to European ash tree pollen [9, 36]. In Locarno and Lugano, Switzerland, 51.8% of pollen allergic patients were found to be sensitized to ash, whereas in Zurich – 56% [37, 38]. A study conducted in Vienna, Austria, revealed that the group of *Fraxinus* pollen-sensitive patients accounted for 17.6% [6], but in Warsaw – 7.7% [39]. Based on the analysis of *Fraxinus* pollen seasons in 7 Polish cities, it was found that the highest number of days with high pollen concentrations was observed in Lublin [34, 35].

The current study shows that in Lublin the *Fraxinus* pollen season begins in April and ends in the first 10 days of May. Because this is simultaneously the birch pollen season, it can be conclude that the pollen seasons of these 2 taxa co-occur. In Poland, many people suffer from allergies caused by the presence of *Betula* pollen [10]. Because it has been demonstrated that cross-reactions to pollen allergens of these tree genera may occur [7], attention should be drawn to this fact since it seems to have been ignored to-date.

![Figure 6](image_url)  
*Figure 6. The duration of the *Fraxinus* pollen season; actual vs. regression model calculated values and predicted value in 2016*

![Figure 7](image_url)  
*Figure 7. Peak date in *Fraxinus* pollen season; actual vs. regression model calculated values and predicted value in 2016*

![Figure 8](image_url)  
*Figure 8. Annual total of *Fraxinus* pollen; actual vs. regression model calculated values and predicted value in 2016*
on the Betula pollen season [20]. Fraxinus pollen season was found to be largely modified by meteorological conditions in February (Tab. 4).

In the presented study, PCA method was used to show in one plot the diversity of individual seasons in terms of the investigated season parameters. A similar technique applied to characterize the studied seasons, this time using meteorological characteristics, was presented in the papers concerning Betula [20] and Poaceae pollen concentrations [21]. Graphical presentation of the results obtained using the PCA method facilitates quick evaluation of Fraxinus pollen seasons that have similar characteristics, and allows us to easily find seasons that exhibit significant deviations from the others.

Regression analysis is frequently used in aerobiological research as a casual method of statistical modeling [5]. The relationship between pollen season parameters and meteorological data has been investigated using regression analysis in many studies. The above-mentioned statistical analyses are used to forecast parameters of the pollen seasons of various taxa. For this purpose, average monthly values of weather data are used [17, 40, 41] or, as in the present study, data from 10-day periods that allowed more accurate regression models to be developed [42, 43, 44].

CONCLUSIONS

Fraxinus pollen grains are present in the air of Lublin usually in April and at the beginning of May. The highest variability was found in the peak value and annual total, while the lowest in the end of the season. There was a strong negative relationship between start and duration and positive between peak value and annual total. The number of days with concentrations above the established threshold was very different in individual years and the highest in 2011 and 2015, while 2016 was characterized by lack of days with high concentrations. The PCA graphical presentation allowed seasons with similar parameters to be identified. The year 2006 proved to be the most outlying season, whereas the seasons in 2008, 2010, 2014 and 2015 as well as in 2007 and 2009 were characterized by similarity in terms of season onset, duration and intensity. In the present study, forecast models were developed for the following season characteristics: start, end, duration, peak date, and total annual pollen sum. The obtained models explain 62–94% of the variation in the investigated characteristics. The best model fit was obtained for season start, whereas the poorest fit for season end. The predicted values for the individual season characteristics in 2016 which were calculated based on a regression equation were close to the actual values. The analysis of the data showed that weather conditions mainly in February are important factors controlling the Fraxinus pollen season.

REFERENCES


