



Predicting the start of the birch pollen season at London, Derby and Cardiff, United Kingdom, using a multiple regression model, based on data from 1987 to 1997

B. Adams-Groom^{1,*}, J. Emberlin¹, J. Corden², W. Millington² & J. Mullins³

¹National Pollen Research Unit, University College Worcester, Henwick Grove, Worcester, WR2 6AJ, UK;

²Midlands Asthma and Allergy Research Association, 12 Vernon St., Derby, DE1 1FT, UK; ³Llandough Hospital, Cardiff, S. Glamorgan, UK

(*author for correspondence: Phone: +44 (0) 1905 855200; Fax: +44 (0) 1905 855234)

Received 17 June 2002; accepted 17 June 2002

Key words: aerobiology, birch pollen season, predictive model, start dates

Abstract

Recent changes in weather in North-West Europe have been reflected in the start dates of pollen seasons. It is therefore necessary to update previous models, such as the one produced by Jones (1995), so that the model will be weighted by current weather patterns. Birch pollen data, collected over a period of eleven years (1987 to 1997 inclusive) from three pollen counting stations in the UK, London, Derby and Cardiff, were analysed to determine the start dates using the Sum75 method. The start dates of the birch pollen seasons of the eleven-year period were then tested for significance against ten-day aggregated variables of temperature and rainfall for each site. The significant variables were entered into multiple regression models until the most valid equation for each site was found. The models were then tested on three years not included in their data sets. The models showed mean differences between actual and predicted start dates, for the eleven years used, of 1.5, 3 and 5 days at Derby, Cardiff and London respectively. For the test years the mean difference was 1, 4.5 and 7.5 days at Derby, Cardiff and London respectively. The most powerful model was for Derby where the corresponding meteorological station is at 0.5 km distance and the weakest was for London where the corresponding meteorological station is much further away at 21 km distance. Weather variables from early February to mid March were found to be the most influential on the start dates of the birch pollen season at the three sites.

1. Introduction

Birch pollen ranks highly as an allergenic species, especially in Northern European countries (Spieksma, 1990), producing symptoms of seasonal allergic rhinitis and asthma in those susceptible. It is useful to be able to forecast when the birch pollen season will begin so that preventative treatment can be started early enough. Information about the start of the birch pollen season may also assist in the planning of clinical trials and is useful for updating pollen calendars.

This paper presents a predictive model using eleven years of birch pollen data from 1987 to 1997.

It is based on the one originally employed by Jones (1995) which used 33 years of pollen and meteorological data from 1961 to 1992. Birch pollen data are used from three pollen sites in the United Kingdom that differ in their climatic regimes and geographical situation, Derby, London (Islington) and Cardiff. Apart from the differences in time period, the methods are the same, using correlation and multiple regression to build a useful model. This updated version focuses on later years to ensure that it is the more recent trends in start dates (Emberlin et al., 1997) and weather patterns that weight the model. Emberlin et al. (1997) found that the recent trend is for birch pollen seasons

to begin earlier. This trend shows a close relationship to an increase in cumulative temperatures in the months of January, February and March recorded near each of the sites during the study period. Jones' model (1995) was accurate in predicting the birch pollen start date of years whose meteorological data was incorporated into it but was very weak when tested on years not included in it. Therefore, the current work aimed to improve the predictive power of the model when testing it on years outside the data set.

The variables influencing the start of the birch pollen season are well known from previous work: Chandler (1925), Kramer and Kozłowski (1979), and Brown (1960) all discussed various aspects of vernalisation in birch trees. In 1989, Faust reported that the actual time of bloom in the spring was dependent on the spring temperatures as warmth is required for bud growth. Jones' model used temperature and rainfall data from several months prior to the start of the birch pollen season. Emberlin et al. (1993) reported that variables such as relative humidity, wind direction and velocity were significant in the prediction of day to day variations once the season had started, but they did not contribute to the prediction of broad seasonal patterns. Spieksma et al. (1995) found that air temperature in the two months preceding the beginning of the birch pollen season was strongly correlated with the start dates. Andersen (1991) used Chill Units and Growing Degree Hours and gained a prediction of within 3–5 days for *Betula* in Denmark for the period 1977 to 1990. Andersen (1991) suggested that while temperature is the main factor, photoperiod might also be significant although it was not included in that model. Other variables such as sunshine hours and synoptic situation are themselves correlated with temperature and rainfall. Norris-Hill (1998) took the start dates of the *Corylus* and *Taxus* pollen seasons and entered them into a multiple regression equation with the mean February temperatures to predict the start of the birch pollen season in London. Models which use Chill Units and Growing Degree Hours (Andersen, 1991) and the start of the pollen seasons of earlier flowering species to predict the start of the birch pollen season (Norris-hill, 1998) seem to be successful at making predictions. However, such data are not available for all pollen counting sites and the models are therefore limited in value. In contrast, temperature and rainfall data usually are available.

Table 1. The distance and direction of the meteorological stations from the pollen monitoring sites.

Meteorological station	London (Heathrow)	Cardiff	Derby
Position	51.48 N 0.48 W	51.42 N 3.2 W	52.93 N 1.5 W
Altitude (m)	25	67	95
Distance from pollen site [km]	21	14.5	0.5
Direction from pollen site	SW	SW	SW

2. Materials and methods

2.1 Data sources

The London pollen data was collected from the University of North London from 1987 until 1995, and then from Islington Council offices where birch pollen counts resumed in 1997, so there is no birch pollen data for 1996. Pollen data for Derby comes from the Midlands Asthma and Allergy Research Association (MAARA). The Asthma and Allergy Unit, Sully Hospital, Penarth, collected the Cardiff Pollen data from the University building near the centre of the city.

The birch pollen in the UK comes from the species *Betula pubescens*, *Betula pendula* and a number of hybrids. The species are not distinguished in the pollen counts. The meteorological data, temperature and rainfall, was used from the stations shown in Table 1. Eleven years of data to 1997 was incorporated into the model and a further three years were used for testing it. No other sites in the United Kingdom have enough years of pollen data to use in both the model and for testing (a minimum of ten years was deemed necessary to provide an adequate sample for the model). The pollen data set used to produce the original model for forecasting the start of the birch pollen season at Cardiff, Derby and London spanned 30 years (early 60s to 1992) (Jones, 1995).

2.2 Start of season

The Sum 75 method of defining the start of the season was used, whereby the start of the season is defined as the first day when the cumulative daily count reaches 75 grains (Jones, 1995). Within and between site variation in the start of the birch pollen season was analysed using ranges and standard deviation.

2.3 Predictive model

The parameters used in the current analysis for predicting the start of the birch pollen season, are mean daily temperatures and total rainfall. Emberlin et al. (1993) showed that meteorological variables aggregated into monthly periods did not produce a particularly high level of explanation. Other authors such as Jones (1995) and Spieksma et al. (1995) used ten day aggregated variables (also known as decades-of-days). The current analysis uses the latter method, calculated 200 days from the mean start date of the birch pollen season. This period includes the weather in the previous spring, winter and autumn. The variables and the start dates were tested for correlation using Pearson's Product. Where linear relationships occurred between the variables and the start of season dates, those variables with a probability of 10% or less were processed through a linear multiple regression analysis. The variables were processed in several combinations until the strongest model for the start of season prediction was constructed.

2.4 Testing the model

Each model was tested on three years of data that were not included in the model, 1982 and 1983 for London, 1985 and 1986 for Cardiff and 1983 and 1985 for Derby, plus 1998 for all three sites. It was not possible to use the same test years for all three sites as Derby and London both have some years pre-1987 in which birch data was not collected.

3. Results

3.1 Within and between site variation in start date

The range of start dates at each site and the standard deviation for each mean start date are given in Table 2. London, in the Southeast of England, has the earliest mean start date for the period at 94 days from January 1st. This is seven days earlier than Derby, which is the most northerly of the three sites, and nine days earlier than Cardiff in the coastal West. The range of start dates spans 26, 25 and 24 days at London, Cardiff and Derby respectively.

3.2 The predictive models

The equations that have been selected as providing the best estimate of the start of the birch pollen season are as follows:

Table 2. Within and between site variation in start dates at London, Cardiff and Derby for the period 1987–1997.

Site	London	Cardiff	Derby
Mean start date as days from January 1st	94	103	101
Standard deviation	8.96	8.28	8.58
Mean start date	4th April	13th April	11th April
Range of start dates as days from January 1st	82–108	89–114	87–111

(i) London

An Rsq value of 75.8% was achieved using the variables shown in the equation:

$$\text{Sum75 Start: } 113.52 - 2.07a - 0.92b + 0.54c - 0.17d$$

- a = Total rainfall 4–13 February
- b = Mean temperature 14–23 February
- c = Mean temperature 24 February–5 March
- d = Mean temperature 6–15 March

Figure 1 shows the predicted and actual values for the start of the birch pollen season at London, 1987 to 1997. The mean number of days difference between predicted and actual start dates for the period is 5 days with a range of 1 to 13 days.

(ii) Cardiff

An Rsq value of 81.4% was achieved using the variables shown in the equation:

$$\text{Sum75 Start: } 131.98 - 2.22a - 1.17b - 0.11c - 0.12d$$

- a = Mean temperature 27 March–5 April
- b = Mean temperature 15–24 February
- c = Total rainfall 17–26 November
- d = Total rainfall 27 November–6 December

Figure 2 shows the predicted and actual values for the start of the birch pollen season at Cardiff, 1987–1997. The mean number of days difference between predicted and actual start dates for the period is 3 days with a range of 1 to 7 days.

(iii) Derby

An Rsq value of 95% was achieved using the variables shown in the equation:

$$\text{Sum75 Start: } 113.5 - 2.1a - 1.11b + 0.28c + 0.24d - 0.1e$$

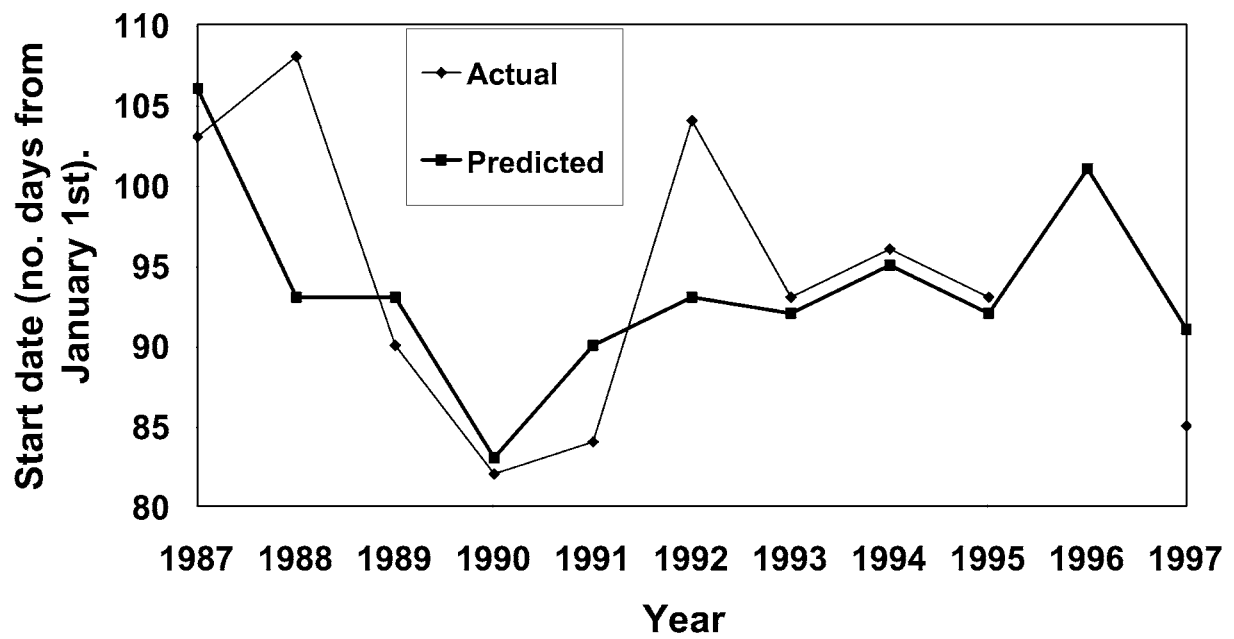


Figure 1. Actual and predicted values for the start of the birch pollen season at London 1987–1997.

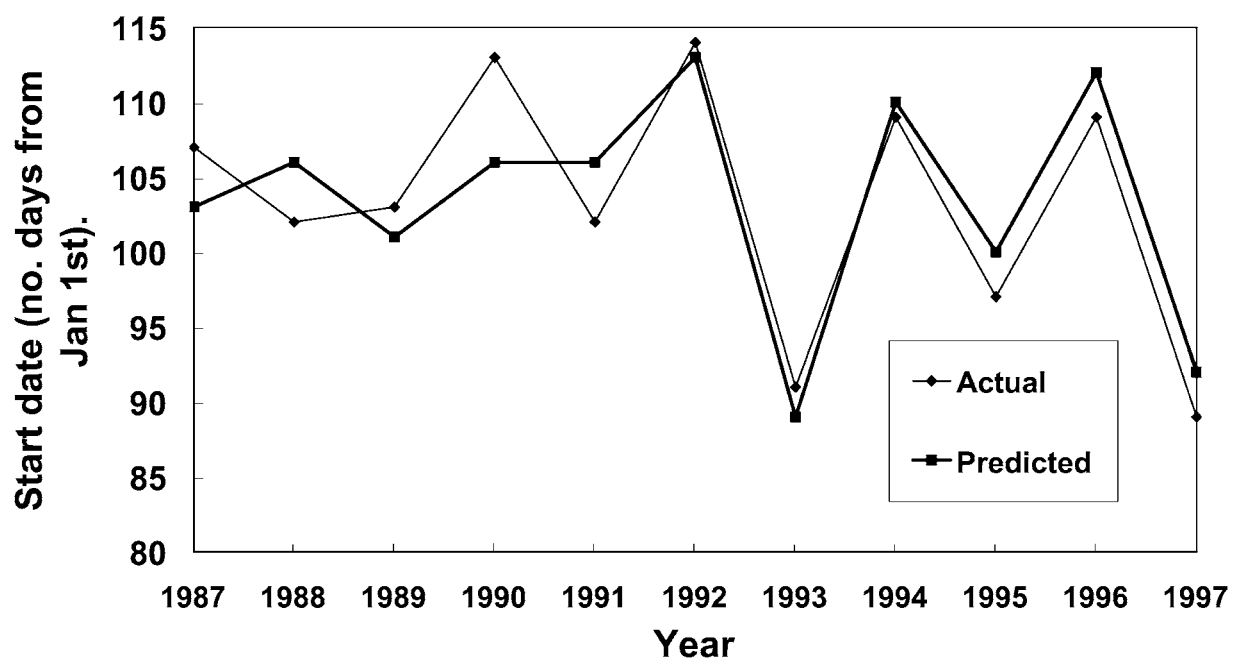


Figure 2. Actual and predicted values for the start of the birch pollen season at Cardiff 1987–1997.

- a = Mean temperature 13–17 March
- b = Mean temperature 16–25 February
- c = Total rainfall 13–17 March
- d = Total rainfall 3–12 March
- e = Mean temperature 3–12 March

Figure 3 shows the predicted and actual values for the start of the birch pollen season at Derby, 1987–1997. The mean number of days difference between predicted and actual start dates for the period is 1.5 days with a range of 0 to 4 days.

3.3 Testing the predictive models

Table 3 shows the results of testing the equations on the London, Cardiff and Derby data respectively. For London the three years had a predicted start date range of 1 to 9 days. Cardiff had a range of 6 to 9 days and Derby 0 to 2 days. London and Cardiff models both give predicted dates that are earlier than the dates actually recorded.

4. Discussion.

All three models show a notable improvement in the current analysis over the original version. This is mainly due to the careful consideration given to which ten-day periods were really affecting the start dates. The Derby version of the model to predict the start of the birch pollen season using the sum 75 method proved to be the most powerful. The Cardiff and London versions of the model have a tendency to predict slightly earlier than the actual start dates, and this is upheld in the tests. It is possible that this occurs because the model uses recent years with mean earlier start dates and it is interesting to note that the test results using 1998 at Cardiff and London are closer to the actual dates than test results using data from the early eighties. Clearly using the test years from before the data set included in the model is not ideal but to have tested only on 1998 would have been inadequate.

Derby had the strongest correlations between weather variables and start dates and was the most straightforward to prepare. Cardiff and London had fewer and weaker significant variables. It is possible that the influence of the weather is a very localised matter when it comes to the start of the birch pollen season at any individual site, e.g. Derby pollen site's nearest meteorological station is only 0.5 km away, while Cardiff and London are much further at 14.5

and 21 km respectively. Therefore, the weather data used in the model may not correspond very closely to those at the trap site, except at Derby. Cardiff is near the sea where coastal effects can produce very different pollen counts to those inland. Initially, it did not seem sensible to include the rainfall variables from late November and December for Cardiff even though they were given as significant in the Pearson Product tests. It is assumed that the trees would be dormant at this time but the rainfall during this period was to prove essential to the creation of a useful model.

London's model was the least accurate of the three, partly because of the climatic differences between pollen and meteorological sites and partly because there are more factors for this site to be taken into account. The majority of pollen collected by a trap will, under usual weather conditions have come from local sources (Jones, 1995). Emberlin and Norris-Hill (1991), however, found that for London the input of birch pollen from remote areas seemed to exceed that from the immediate vicinity where there were few birch trees. Work by Emberlin et al. (1993) and Jones (1995) has shown that the start dates of the pollen seasons are more closely related to weather in the pollen source areas than in central London where the influence of the urban climate predominates. Buildings can modify the local airflow and increase turbulence which affects pollen dispersal and they can also affect pollen concentrations through impaction and filtration on their surfaces (Emberlin and Norris-Hill, 1991). Given that there are so many different variables affecting the London pollen data, it would seem that obtaining a mean difference between predicted and actual start dates of only five days is a very acceptable result.

Although the ten-day aggregated variables were considered for the previous two hundred days to the previous autumn of the year before, it was mainly the later decades of days that were found to be of use in the model. This indicates that the factors involved in the initial birch flower bud development in the autumn are minimal in affecting the start of pollen release the following spring. The variables that were most significant were not necessarily the most important for the creation of a useful model. For example, at Cardiff, the rainfall variable for mid-December was found to correlate significantly with the start dates but including it produced a model with poor predictive powers. With the exception of Cardiff, the models' variables derive from February through to 20 days prior to the mean start date (Cardiff has some late autumn variables and

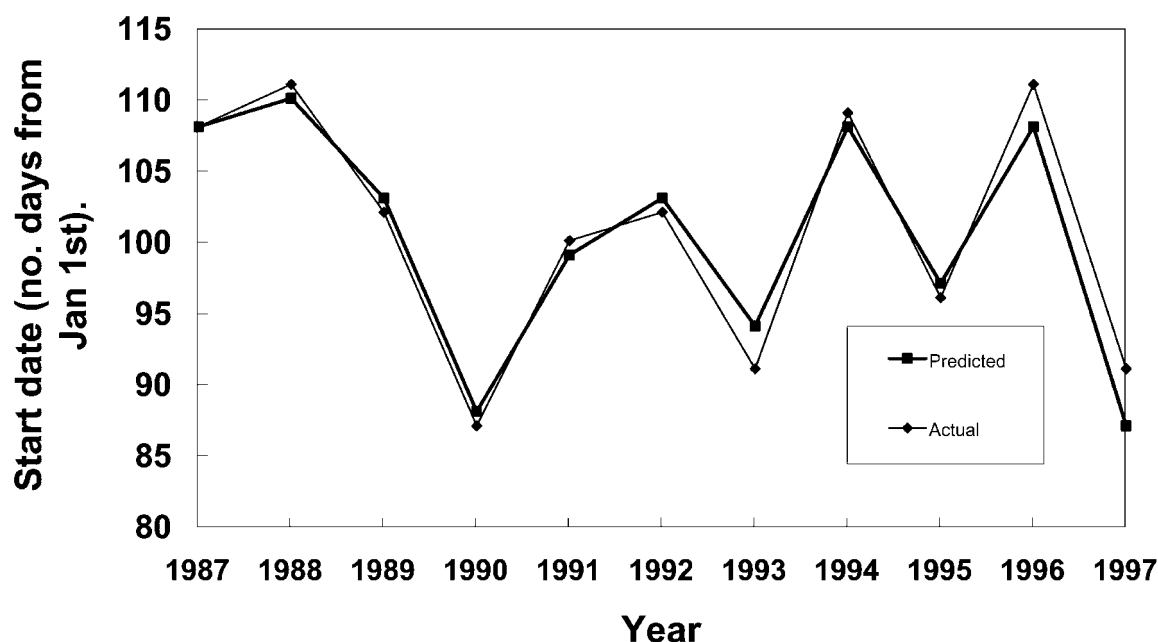


Figure 3. Actual and predicted values for the start of the birch pollen season at Derby 1987–1997.

Table 3. Test results for the model to predict the start of the birch pollen season at London, Cardiff and Derby. An earlier prediction is denoted by + and later by –.

Test year	London			Cardiff			Derby		
	1982	1983	1998	1985	1986	1998	1983	1985	1998
Predicted start	98	96	93	98	113	88	107	110	93
Actual start	112	105	92	107	121	94	107	108	93
No. days difference	+4	+9	–1	+9	+8	+6	0	–2	0

data through to 10 days prior to the mean start date). The variables covering the interim period (i.e. the 20 days prior to the mean start date) at London and Derby were only very weakly correlated with start dates and led to reduction in the power of the models when included. A spell of extremely cold, wet or unusually warm weather occurring in this period could delay or hasten the opening of the birch tree anthers and would therefore lie beyond the power of the model. Examination of the weather data used seems to suggest that this is indeed the case. For example, at London (Heathrow) in 1988 there was very heavy rainfall during this interim period and the actual start date was 13 days behind the predicted start, while the reverse happened in 1997 when very dry and mild weather produced an earlier start than predicted. It is suggested that a long-

range weather forecast be used in conjunction with the model to cover this time period.

Acknowledgements

M. Savage, St. Mary's Hospital Trust, Paddington, London, W2 1PG, UK, for London pollen data 1982 and 1987. Prof. R. Eccles, Common Cold Centre, Cardiff University, Cardiff CF1 3US, for 1997 and 1998 Cardiff pollen data.

References

Andersen T.B.: 1991, A model to predict the beginning of the pollen season. *Grana* **30**, 269–275.

- Brown D.S.: 1960, The relation of temperature to the growth of apricot flower buds. *Proc Amer Soc Hort Sci* **75**, 138–147.
- Chandler W.H.: 1925, *Fruit Growing*. Houghton Mifflin, Boston.
- Emberlin J. and Norris-Hill J.: 1991, Spatial variation of pollen deposition in North London. *Grana* **30**, 190–195.
- Emberlin J., Savage M. and Woodman R.: 1993, Annual variations in the concentrations of *Betula* pollen in the London area, 1961–1990. *Grana* **32**, 359–363.
- Emberlin J., Mullins J., Corden J., Millington W., Brooke M., Savage M. and Jones S.: 1997, The trend to earlier birch pollen seasons in the UK: A biotic response to changes in the weather conditions? *Grana* **36**, 29–33.
- Faust M.: 1989, *Physiology of Temperate Zone Fruit Trees*. John Wiley and Sons, London, pp. 169–234.
- Jones S.: 1995, Allergenic pollen concentrations in the United Kingdom. PhD thesis, University of North London.
- Kramer P.J. and Kozlowski T.T.: 1979, Environmental control of reproductive growth. *Physiology of Woody Plants*, 688–693.
- Norris-Hill J.: 1998, A method to forecast the start of the *Betula*, *Platanus* and *Quercus* pollen seasons in North London. *Aerobiologia* **14**(2,3), 165–170.
- Spieksma F.Th.M., Emberlin J.C., Hjelmroos M., Jager S. and Leuschner R.M.: 1995, Atmospheric birch (*Betula*) pollen in Europe: Trends and fluctuations in annual quantities and the starting dates of the seasons. *Grana* **34**, 51–57.

