ELSEVIER

Aerobiologia 13 (1997) 177–184

Aerobiologia
International Journal of Aerobiology

Using meteorologic data to predict daily ragweed pollen levels

Paul C. Stark [a,*], Louise M. Ryan [a,b], James L. McDonald [c], Harriet A. Burge [d]

[a] *Department of Biostatistics, Harvard School of Public Health, 677 Huntington Avenue, Boston, MA 02115, USA*
[b] *Dana-Farber Cancer Institute, Boston, MA 02115, USA*
[c] *2490 South 11th Street, Kalamazoo, MI 49009, USA*
[d] *Department of Environmental Health, Harvard School of Public Health, 665 Huntington Avenue, Boston, MA 02115, USA*

Received 7 January 1997; accepted 11 June 1997

Abstract

Pollen-related allergy is a common disease resulting in symptoms of hay fever and asthma. Control of symptoms depends (generally) on avoidance and pharmacological treatment. Both of these approaches could benefit from accurate predictions of pollen levels for future days. We have constructed a model that uses meteorological data to predict ragweed pollen levels based on air samples collected daily in Kalamazoo, MI from 1991 to 1994. Ragweed pollen counts were converted to pollen grains/m^3 of air (24-h average). We used Poisson regression, which appropriately handles the heterogeneous variance associated with pollen data. Using standard statistical model selection procedures, combined with biological considerations, we selected rainfall, wind speed, temperature, and the time measured from the start of the season as the most significant variables. Using our model, we propose a method that uses the weather forecast for the following day to predict the ragweed pollen level. This approach differs from most previous attempts because it uses Poisson regression and because this model needs to be fit iteratively each day. By updating the coefficients of the model based on the information to date, this method allows the fundamental shape of the pollen distribution curve to change from year to year. Application to the Kalamazoo data suggests that the method has good sensitivity and specificity for predicting high pollen days.

Keywords: Ragweed pollen; Poisson regression; Forecasting; Meteorology

1. Introduction

Pollen and fungal spores are the most abundant biological particles in the atmosphere (Arizmendi et al., 1993). Pollen-related allergy is a common disease resulting in symptoms of hay fever and asthma in about 10% of the population. Control of pollen-induced symptoms depends (generally) on avoidance and pharmacological treatment. Both of these approaches could benefit from accurate predictions of pollen levels for future days.

The relationship between pollen levels and meteorological variables has been clearly established. O'Rourke (1988) used a χ^2 goodness-of-fit test to verify the existence of a positive relationship between temperature and pollen dispersal. Temperature, rainfall and wind have been shown to be significant predictors of ragweed and oak pollen levels using multiple linear regression (Reiss and Kostic, 1976; Fairley and Batchelder, 1986). McDonald (1980) used correlation coefficients to support the importance of wind in grass pollen dispersal. Correlation coefficients have also been used to demonstrate that there is a strong relationship between these weather variables and 'weed', birch and grass pollen concentrations (Bringfelt et al., 1982; Glassheim et al., 1995). In addition, Frenz et al. (1995) establish that latitude influences the date of maximum ragweed pollen concentration. However, latitude may simply be a surrogate for temperature in their studies.

* Corresponding author. Tel.: +1 617 4321056; fax: +1 617 7391781. Also affiliated with Massachusetts General Hospital, Boston, MA 02114, USA; Tel.: +1 617 7246985; fax: +1 617 7244015.

0393-5965/97/$17.00
PII S0393-5965(97)00025-5

While the relationship between these weather variables and pollen production is undisputed, use of statistical models to predict the daily ragweed pollen counts have been less successful (Goldberg et al., 1988) possibly due, in part, to inadequacies in the models. A more recent attempt, discussed by Comtois and Sherknies (1992), develops a forecasting model based on the probabilistic distribution of ragweed pollen curves; we take a different approach. We have focused our efforts on predicting pollen levels for the following day on ragweed (*Ambrosia* spp.) pollen from a single Midwestern site (Kalamazoo, MI) and have used environmental variables and Poisson regression to develop our model. It seemed reasonable to focus our analysis on ragweed pollen because pollen derived from *Ambrosia* spp. has long been recognized as a major cause of pollinosis (Buck and Levetin, 1983). Also, ragweed pollen in the Midwest is shed by only two species (*A. artemisiifolia* and *A. trifida*). In addition, the factors that control the production of ragweed pollen are consistent and well-documented. Ragweed initiates pollen production only when periods of darkness (nighttime) are sufficiently long. Thus, in temperate regions, ragweed pollen is usually not available for dispersal until late July (when nights become long enough). In addition, for ragweed, the overnight temperature must exceed about 10–15°C (50–60°F) for anther extension to occur (Bianchi, 1959). Therefore, little additional pollen is released after the beginning of October (after the first frost) in Northern climates.

The goal of this work is twofold. Firstly, we develop a model, using the framework of Poisson regression, to determine which environmental variables have the most significant impact on pollen counts. These will be the independent variables in the model, with pollen counts, in grains/m^3 of air (24-h average), as the dependent variable. Next, we utilize this model and the weather data (since many of the explanatory variables are weather-related) in an attempt to predict the ragweed pollen levels for the following day.

2. Materials and methods

2.1. Field data

2.1.1. Pollen collection

Pollen was collected 7 days a week from 1991–1994 using a Rotorod Aeroallergen Model sampler (Burge and Solomon, 1987) located on a flat roof of the local television station about 20 feet above grade. The station is in a suburban neighborhood 1.25 miles south of the city center. Just to the northwest is an 160 acre nature preserve surrounding a small eutrophic lake.

The Rotorod is a rotating arm impactor, rotating silicone-greased plastic I-rods intermittently 1 min in 10 at 2400 rpm. After each 24-h exposure period, between 17.00 and 19.00, rods were removed and one of each pair mounted in Calberla's solution and ragweed pollen counted on the entire rod surface at 600 × using an A.O. 150 microscope. Counts were converted to grains/m^3 of air (24-h average). We chose Kalamazoo because several years of consistently collected data were available and the site generally reported ragweed pollen levels above the overall average for the American Academy of Allergy, Asthma and Immunology's aeroallergen network (AAAAI, 1995).

2.1.2. Meteorological data collection

Meteorological data, collected at the Kalamazoo airport, 3 miles southeast of the pollen collection site, were obtained from the National Center for Atmospheric Research.

2.2. Statistical packages

Data manipulation and the creation of variables occurred in SAS (SAS Institute, Cary, NC). However, the data analysis was performed in **S-PLUS** (Seattle: Statistical Sciences, a division of MathSoft Inc., Version 3.4).

3. Development of the model

3.1. Motivation for using Poisson regression

Scatterplots of pollen levels over time indicated that the mean pollen levels follow an inverted U-shaped curve, with a slightly skewed right tail (Comtois and Sherknies, 1987) (Fig. 1). It is clear from Fig. 1 that as the mean pollen level increases, so does the variation of the pollen levels. This property suggests that Poisson regression may be appropriate since this technique assumes that the variance of the data is proportional to its mean. Poisson regression is often the appropriate statistical model to use for data that will always take on integer values, as in our case. Poisson regression models the mean (or expected value, denoted EY) of the response as a function of covariates on the natural logarithm scale. For example, if X_1 and X_2 are two covariates of interest, and Y is the outcome, Poisson regression assumes that:

$$\log(EY) = \alpha_0 + \alpha_1 X_1 + \alpha_2 X_2$$

or,

$$EY = \exp(\alpha_0 + \alpha_1 X_1 + \alpha_2 X_2).$$

More specifics about the predictor variables to be considered in the model will be given presently. In this paradigm, all models will include the variables **day** and ln(**day**). Ignoring other covariates for now, this implies that the mean pollen levels will have the following form:

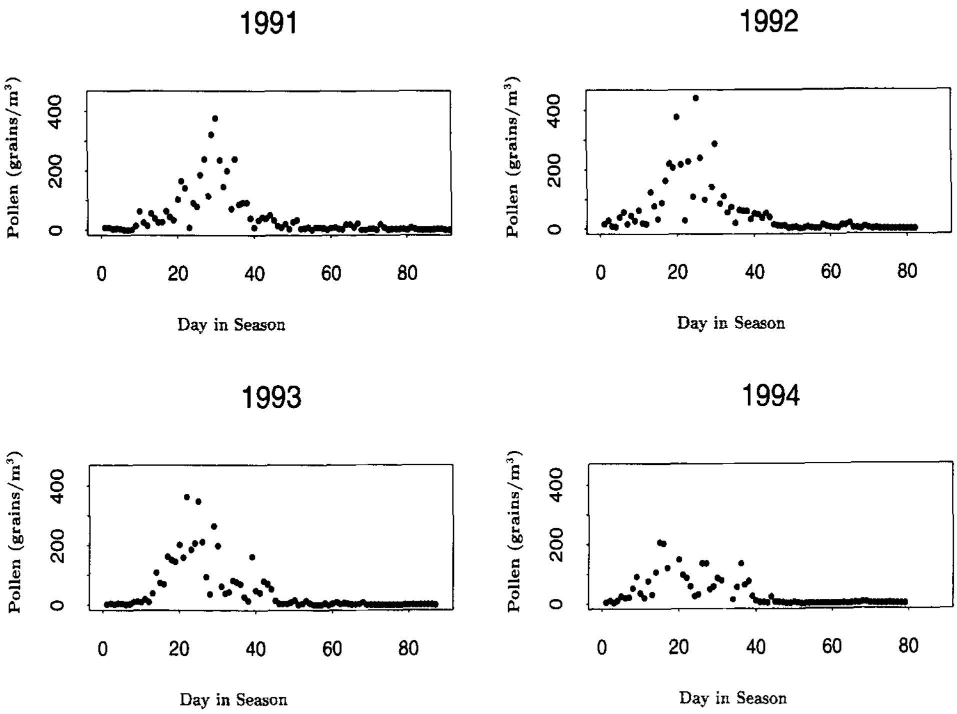

Fig. 1. Pollen grains/m³ of air (24-h average), over time.

$$EY = \exp(\alpha_0 + \alpha_1 \text{day} + \alpha_2 \ln(\text{day}))$$

or,

$$EY = \exp(\alpha_0)\exp(\alpha_1 \text{day})\text{day}^{\alpha_2}$$

This function has the required flexibility to allow the mean pollen counts to rise, reach a peak, then fall off again to zero. Poisson models were fit by maximum likelihood using the **glm** function in **S-PLUS**. To adjust for overdispersion, all standard errors were multiplied by a scaling estimate that was estimated as the scaled deviation divided by the degrees of freedom (McCullagh and Nelder, 1992, pp. 199). All reported P-values are based on 2-sided Wald tests.

3.2. Creation of explanatory variables

Aerobiologists generally agree that the day-to-day variation in the concentration of airborne pollen is likely to be related to daily air temperatures, wind velocity, the amount and duration of precipitation and the water content, (i.e. humidity) of the air (Moseholm et al., 1987) The meteorologic data available to us included hourly observations of a myriad of weather-related factors. These observations were summarized into daily values.

The primary time period of pollen dispersal is during the late morning, so it is important to note whether or not there was significant rain during this period. Hence, we created the binary or indicator variable: **rain** =

- 0 if there were at least 3 h of steady rain or brief but intense rain (in the late morning),
- 1 otherwise.

Similarly, an overnight temperature below 50°F (10°C) is likely to decrease pollen dispersal for the following day. We therefore created the indicator variable: cold =

- 0 if the overnight temperature ever dropped below 50°F,
- 1 otherwise.

Wind speed in knots (**wind**) is also included in the analysis as a continuous or quantitative variable since it has been shown that wind tends to influence pollen dispersal in a linear manner (McDonald, 1980).

In addition, functions of temperature (in °F) need to be considered. Daily average temperature was an informative variable. However, having only daily average temperature in the model will not fully explain the effect of temperature. For example, an 80°F (27°C) day may have a much different influence on pollen production in late September than it will in late July. Instead, two functions of temperature were created. They were a smooth function of temperature (**temp.trend**), which was created using the **loess** function in **S-PLUS**, and the deviation of the daily average temperature from the **loess** line (**temp.resid**) (Fig. 2). The former can be thought of as a smooth function of the average temperature, while the latter as the departure from the average expected temperature on any given day. Akaike Information Criterion (AIC) was used to determine the

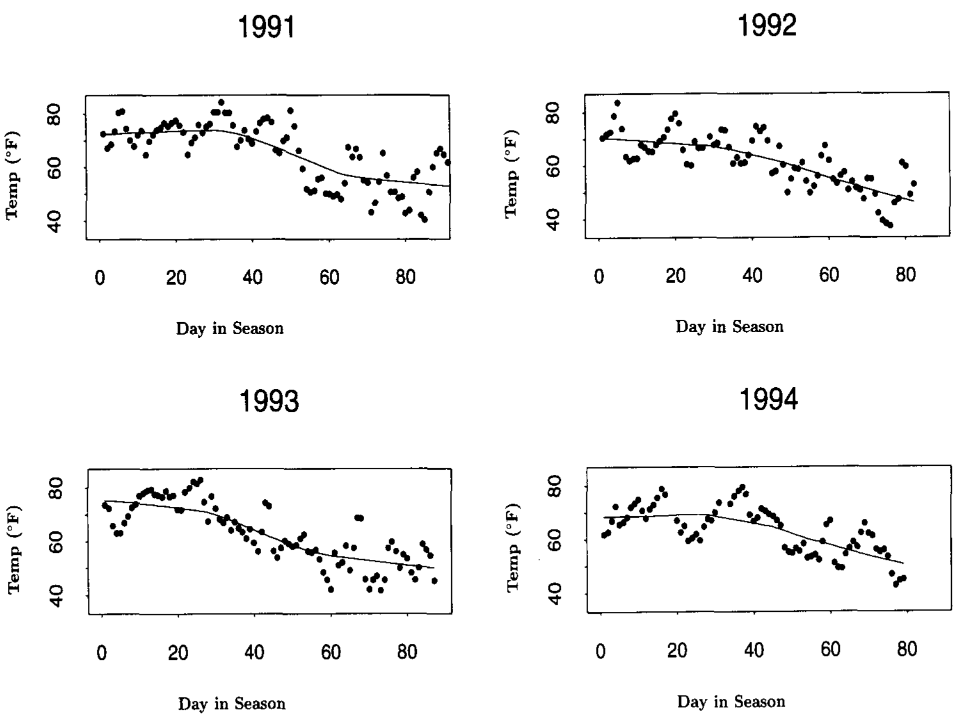

Fig. 2. Average daily temperature (in °F), over time.

optimal window size for the **loess** functions (McCullagh and Nelder, 1992, pp. 91).

Temp.trend and **temp.resid** allow the model to distinguish between the two previously described temperature scenarios. Thus, an 80°F day in September is likely to be associated with a high value for **temp.resid** and a low value for **temp.trend**, whereas the same type of day in late July will have the opposite values for the corresponding parameters.

Finally, we needed a means of defining the appropriate time frame for our analyses. We created a variable called **day** that is, simply, the day in the season, with day 1 being the first of 4 consecutive non-zero pollen days. For example, for 1991, day equals 1 on the 209th day of the calendar year (28 July), 2 on the 210th day of the calendar year (29 July), etc... Justification of this choice for the start date criterion will be giving in the Section 5. Since **day** will clearly be a monotonically increasing variable (i.e. it will always increase, never decrease), and the pollen levels start to decrease midway through the season, we needed an additional function of **day** to accommodate this nonlinear pattern. As discussed in the previous section, the natural logarithm of the day in the season, ln(**day**), was useful for this purpose. The inclusion of **day** and ln(**day**) as predictors in the Poisson regression provides a simple, yet flexible modeling framework that can easily capture the typical pattern of pollen dispersal over the season.

4. Results

4.1. Fitting the model

The primary objective in the model selection process was to determine a representation of the biologically meaningful variables that provided a parsimonious model (i.e. one with only a few variables) that would best allow for the prediction of the pollen levels. After a series of stepwise regression procedures and examination of the resulting residuals to assess the model fit (Fig. 3), we arrived at a model that meets these requirements. The model included the following variables: **rain** (the indicator variable), **day**, ln(**day**), **temp.trend**, **temp.resid** and **wind**. Because we are using Poisson regression, the expected pollen level on any given day takes the form:

$$E[\text{pollen}] = exp(\alpha_0 + \alpha_1 X_1 + \alpha_2 X_2 + \alpha_3 X_3 + \alpha_4 X_4 + \alpha_5 X_5 + \alpha_6 X_6),$$

where the α terms are the unknown regression coefficients to be estimated and the X's are values of the corresponding variables (i.e. **day**, ln(**day**), etc.). Contrary to expectations, **relative humidity** and **cold** (the indicator variable) did not appreciably improve the model, and were ultimately not included. The influence of **cold** may already be explained by the inclusion of the two temperature variables, and the primary effect of **relative humidity** by the **rain** variable.

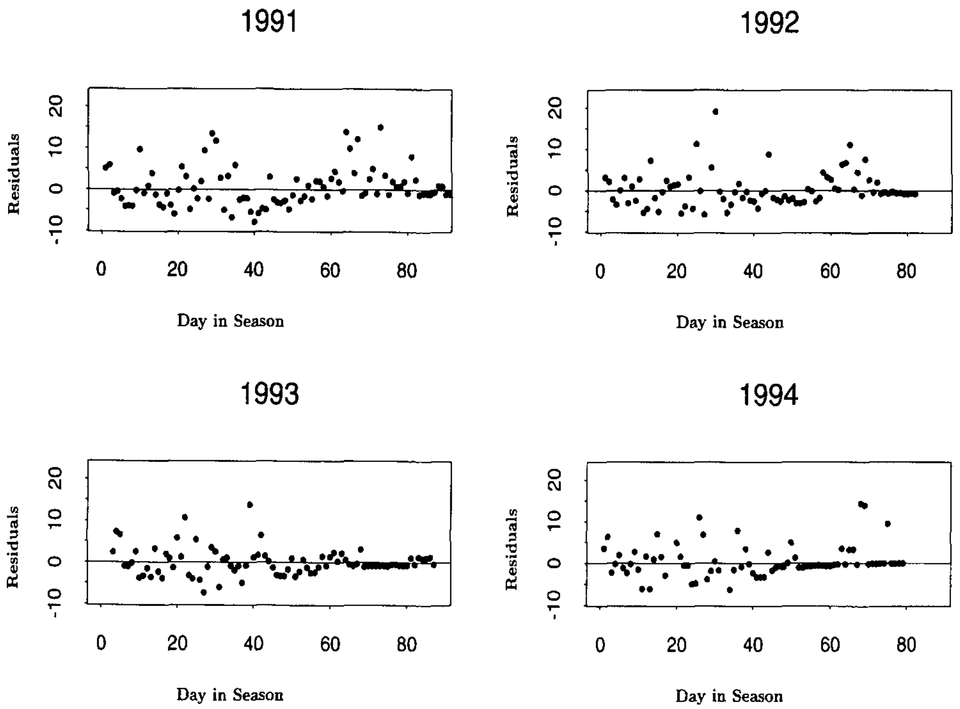

Fig. 3. Pearson residuals of the pollen prediction model, over time.

Table 1 summarizes the final model fits for all 4 years. Ideally, the sign, magnitude and significance of the estimated coefficient of any given variable should be consistent across the 4 years, especially if the coefficients from the final model from 1 year were going to be used to predict the pollen for the following year. As can be seen from Table 1, the consistency holds reasonably well. In general, inspecting the estimated coefficients in Table 1 allows us to assess the influence of each individual variable on the predicted pollen level, while holding all other variables constant. The one exception to this is the pair of variables **day** and ln(**day**). It does not make sense to try to interpret these separately, but rather, the pair together allows the mean model to assume its basic shape: that is, pollen levels start at zero, rise during the main pollen season, then decline back to zero. The purpose of additionally modeling the meteorological variables is to account for day-to-day departures from this basic shape.

4.2. The prediction model

Having established a model that explains the observed patterns in pollen dispersal, our next goal was to use this model to develop a strategy for prediction based on the weather forecast for that day. Clearly, there will be some imprecision in the forecasts. It becomes especially difficult to forecast the indicator variable, **rain**. We suggest that the **rain** variable be assigned the value 0 if the weather forecast calls for a chance of late morning rain for the following day to be greater than or equal to 50%.

The proposed algorithm for predicting ragweed pollen levels is as follows. Note that one must have 7 days of data (the 4 consecutive non-zero days, plus 3 additional days) before a prediction can be made.

1. Start gathering data for the variables identified in Section 4.1, and begin to collect the pollen counts.
2. Determine the start of the season (assign **day** = 1 for the first of 4 consecutive non-zero pollen days).
3. Starting at **day** = 7, create the derived variables (**temp.trend**, **temp.resid** and **rain**) and run a Poisson regression on the data, where pollen count is the predicted variable (there will be seven observations for each variable).
4. Check to see if the coefficient for ln(**day**) is significant. If it is not, drop it from the model and rerun.
5. Use the weather forecast for the following day and the coefficients from the model to calculate the predicted pollen for the following day. This is accomplished by using the fact that:

$$E[\text{pollen}] = \exp(\alpha_0 + \alpha_1 X_1 + \alpha_2 X_2 + \alpha_3 X_3 + \alpha_4 X_4 + \alpha_5 X_5 + \alpha_6 X_6)$$

6. On the next day, update the data based on the actual observed values, and return to 3.

Table 1
Final parameter estimates from prediction

Parameter	1991			1992			1993			1994		
	Value	S.E.	*P*-value	Value	S.E.	*P*-value	Value	S.E.	*P*-value	Value	S.E.	*P*-value
Intercept	−25.579	3.837	<0.001	−19.751	7.740	0.010	−23.295	4.514	<0.001	−24.710	6.325	<0.001
Rain	0.768	0.457	0.093	0.997	0.451	0.031	1.055	0.242	<0.001	0.938	0.417	0.025
Day	−0.002	0.039	0.965	−0.101	0.043	0.021	−0.123	0.049	0.012	−0.104	0.035	0.003
ln(day)	1.393	0.756	0.064	2.771	0.523	<0.001	4.474	0.707	<0.001	2.069	0.593	<0.001
Temp.trend	0.335	0.064	<0.001	0.245	0.113	0.035	0.213	0.069	<0.001	0.339	0.100	<0.001
Temp.resid	0.042	0.015	0.006	0.044	0.011	<0.001	0.040	0.010	<0.001	0.052	0.011	<0.001
Wind	0.057	0.024	0.019	0.095	0.019	<0.001	0.098	0.018	0.002	0.087	0.021	<0.001

To test our prediction model, we applied it to the data from Kalamazoo. Because the weather forecasts were unavailable, we used the actual weather data to predict daily pollen levels using the algorithm described above. By comparing predicted and actual pollen levels, we were able to assess the accuracy of our approach. It is important to note that we have not fallen into the statistical trap of assessing prediction accuracy with the same data used to construct the model. Our approach is different because the prediction process involves refitting the model using only data up to the day to be predicted. Our initial analysis of the entire data set served only to identify the variables to be used in the prediction process.

The iterative method introduced here did usually predict pollen levels that were close to the observed levels. Occasionally, however, a value outside the realm of possible values was predicted. The reason for this is that when running the regression early in the season (when there are so few data points), the coefficients of the variables are prone to experience great changes as new data is added. This problem disappears if the model is used only to predict whether the pollen levels will be *low*, *moderate* or *high*. We used the cutoffs established by the American Academy of Allergy and Immunology specifically for Kalamazoo (AAAI, 1994) and collapsed the *high* and *very high* category into one *high* category, based on the assumption that any pollen count at or above 72 grains/m^3 will result in pollen-related symptoms.

We then created a table that demonstrates the success of our model in predicting the pollen level correctly (Table 2). For example, in 1991 there were 52 low pollen days observed. Our model predicted correctly on 50 of those days. On the other 2 days our model predicted that the pollen level would be *high*. Clearly, one would hope that the majority of the values in the table would be on the diagonal (indicating a correct prediction), which is were the greatest numbers are usually found. Using the AAAI categories, our model works best for predicting *low* or *high* levels.

5. Discussion

5.1. General discussion

While much research has been conducted to understand the factors that influence pollen levels, there has been very little work in developing a model that can adapt to the changes in the weather as the season progresses. Predictive models have recently been reported for grass pollen (Antepara et al., 1995; Norris-Hill, 1995). However, the biology behind the development and dispersal of grass pollen differs significantly from that of ragweed pollen and there is no reason to assume that a grass-based method would work well for ragweed pollen.

Table 2
Observed vs. predicted pollen levels 1991–1994

Predicted	Observed		
	Low (<18)	Moderate (18–71)	High (72+)
1991			
Low (<18)	**50**	10	0
Moderate (18–71)	0	**9**	3
High (72+)	2	3	**15**
Total observed	52	22	18
1992			
Low (<18)	**39**	4	1
Moderate (18–71)	3	**9**	0
High (72+)	0	8	**17**
Total observed	42	21	18
1993			
Low (<18)	**50**	2	0
Moderate (18–71)	5	**6**	4
High (72+)	0	5	**15**
Total observed	55	13	19
1994			
Low (<18)	**43**	5	1
Moderate (18–71)	1	**6**	6
High (72+)	0	5	**8**
Total observed	44	16	15

The most successful work to date addressed at modeling ragweed pollen levels has been in formulating models that fit well after the season had ended. For example, regression analyses by Bringfelt et al. (1982) and Comtois and Sherknies (1987) have resulted in as much as 80% of the variation in ragweed pollen levels being explained by the model. While impressive, these models will not help pollinosis sufferers, since the effects of the ragweed have already been felt before the analysis is completed.

Our model is fit iteratively using all information to date within the season, (i.e. it is meant to be used for prediction only from day-to-day within each ragweed season). An attempt to fit this model at the end of the season and then use the coefficients from that model to predict the pollen levels for the following year was unsuccessful. This is not surprising when one considers all of the factors that affect the amount of pollen that is available for dispersal, virtually none of which are included in this model. For example, the sum of all the daily pollen counts in grains/m^3 for 1993 and 1994 were 4087 and 2422, respectively. Clearly, any attempt to use the coefficients from the final 1993 model to predict the 1994 pollen levels will result in a systematic overestimation. This further justifies why it is not necessary to use some data to develop the model and other data to test it.

The choice of defining the start of the pollen season as the first of 4 consecutive non-zero pollen days was somewhat arbitrary. In a recent study of grass pollen allergy, the start date was defined as the first day in the season at which 100% of a sample of patients under study suffer from symptoms (Antepara et al., 1995. Clearly, this approach will often not be practical. Furthermore, it is not likely to be useful for predicting the start date for a specific pollen type, such as ragweed. Stix and Ferretti (1974) defined the start of each pollen season as the first day of the year on which the cumulative pollen count exceeds 2% of the average yearly counts of the same taxon over the last 35 years. A disadvantage of this method is that it requires a great deal of historical data to implement. The 1994 AAAI report uses the day on which 1% of the total ragweed pollen for the season has been collected (AAAI, 1994). However, this method cannot be used until after the end of the season. Our approach consistently yielded the same start date for our 4 years of data as would have been obtained using 3 or 5 consecutive non-zero days. Our method produced a start date that was within 1 day of the 1% method for 1992 and 1994, but that was 8 days earlier for both 1991 and 1993.

Because this model includes the two functions of time, **day** and ln(**day**), predicted pollen levels automatically go to 0 as the season draws to a close. Thus, for the purpose of fitting our model, it is not necessary to specify a rule for predicting the end of the pollen season. The models represented in Table 1 used all available data on pollen counts, with each season lasting about 80 days. We considered applying some rules to indicate the end of the season. For example, the first of 4 consecutive days with no detectable pollen (pollen counts = 0) as the first day after the season has ended. However, our data suggested that such a rule would result in the omission of several potentially important days. Take 1994, for example. Even a criterion of 11 consecutive days with a pollen count equal to 0 would have resulted in the omission of at least 6 days, some with counts as high as 4 grains/m^3 of air. These counts may not seem that high when compared to the 250 grains/m^3 of air experienced in the middle of that season, but they may be high enough to cause symptoms at the end of the season.

Developing a tool to correctly predict the end of the pollen season is of little biological importance. The choice of the last day of the pollen season will also have very little influence on our model. We implemented the rigid criterion of 4 consecutive days with no detectable pollen to see how our results would change. For the most part, the parameter estimates in Table 1 experience a negligible change (usually on the order of 1%). The only cell in Table 2 that will be affected is the *low-low* cell. On average, this number will be reduced by 10.

5.2. Limitations of the model and future work

One minor limitation of our approach is that the model cannot be used during the first 7 days of the season. The reason for this is two-fold: firstly, to run the regression, there need to be more observations than explanatory variables; and secondly, to calculate the **loess** line for temperature (**temp.trend**), there need to be at least seven data points. However, because at least 4 consecutive non-zero pollen days are required to signal the start of the season, the model will be applicable after just 3 more days. Also, there are rarely dangerously high levels of ragweed pollen during the first week of the season, so this limitation is relatively insignificant.

One final limitation of this approach is that it relies on the weather forecast for the following day. There is inherent imprecision in the weather forecast, and our method must utilize this imperfect tool. Future work in this area will include testing the model using the weather forecast (and not the observed weather) for a given day to make the pollen prediction.

Much current work in the area of forecasting pollen has focused on using pollen data from previous days to help predict current pollen levels. One danger in this is that there is clearly high autocorrelation between day-to-day pollen levels, so one would need to use sophisticated time-series techniques to adjust for this lack of independence. An example of a time-series analysis on grass pollen data can be found in Moseholm et al. (1987). It is unclear if an approach such as this would provide a significantly better prediction than our iterative method. A trade-off exists between the ease of applicability of the model and the exactness of the prediction. The model proposed here can be implemented by anyone with access to the basic meteorologic data, daily pollen levels, and virtually any statistical package. However, extensions of our model to incorporate predictions based on the previous day's pollen level would be useful.

Finally, there is reason to be optimistic that an adaptation of this model could be extended to sites other than Kalamazoo, MI, provided that the locale in question has similar characteristics with regard to the factors that influence ragweed pollen dispersal. The model is most likely to be useful where few closely-related species shed pollen of the ragweed type. In areas where species of *Franceria* contribute significantly to ragweed pollen counts, the characteristic Poisson distribution of pollen levels during each season, which is basic to our model, may not be consistently present. However; the model may be adaptable for other pollen types for which levels consistently follow the Poisson distribution.

Acknowledgements

This work was supported in part by grants from: American Academy of Allergy, Asthma and Immunology, Milwaukee Wisconsin; National Institute of Health Grants: CA 48061 and ES07142.

References

Aeroallergen Monitoring Network, Pollen and Spore Report, American Academy of Allergy, Asthma and Immunology, 1995.

Aeroallergen Monitoring Network, Pollen and Spore Report, American Academy of Allergy and Immunology, 1994.

Antepara I, Fernandez JC, Gamboa P, Jauregui I, Miguel F. Pollen allergy in the Bilbao area (European atlantic seaboard climate): pollination forecasting methods. Clin Exp Allergy 1995;25:133–40.

Arizmendi CM, Sanchez JR, Ramos NE, Ramos GI. Time series predictions with neural nets: application to airborne pollen forecasting. Int J Biometeorol 1993;37:139–44.

Bianchi DE. Pollen release in the common ragweed (*Ambrosia artemisiifolia*). Botanical Gazette 1959;120(4):235.

Bringfelt B, Engstrom I, Nilsson S. An evaluation of some models to predict airborne pollen concentrations from meteorological conditions in Stockholm, Sweden. Grana 1982;21:59–64.

Buck P, Levetin E. Weather patterns and ragweed pollen production in Tulsa, Oklahoma. Ann Allergy 1983;49:272–5.

Burge HA, Solomon WR. Sampling and analysis of biological aerosols. Atmos Environ 1987;21(2):451–6.

Comtois P, Sherknies D. Le pollen de l'Ambroisie (*Ambrosia artemisiifolia* L.): previsions et prevention. Allerg Immunol 1992;24(1):22–6.

Comtois P, Sherknies D. An aerobiological model for pollen forecasting, editors: Comtois P, Sherknies D. 18th Conference on Agricultural and Forecast Meteorology, 8th Conference on Biometeorology and Aerobiology, W. Lafayette, Indiana, American Meteorological Society, Boston, Massachusetts, 1987.

Fairley D, Batchelder GL. A study of oak-pollen and phenology in northern California: prediction of annual variation in pollen counts based on geographic and meteorologic factors. J Allergy Clin Immunol 1986;78(2):300–7.

Frenz DA, Palmer MA, Hokanson JM, Scamehorn RT. Seasonal characteristics of ragweed pollen dispersal in the United States. Ann Allergy, Asthma Immunol 1995;75:417–22.

Glassheim JW, Ledoux RA, Vaughan TR. Analysis of meteorologic variables and seasonal aeroallergen pollen counts in Denver, Colorado. Ann Allergy, Asthma Immunol 1995;75:149–56.

Goldberg C, Buch H, Moseholm L, Weeke ER. Airborne pollen records in Denmark, 1977–1986. Grana 1988;27:209–18.

McCullagh P, Nelder JA, Generalized Linear Models. London: Chapman and Hall, 1992.

McDonald MS. Correlation of air-borne grass pollen levels with meteorological data. Grana 1980;19:53–6.

Moseholm L, Weeke ER, Petersen BN. Forecast of pollen concentrations of Poaceae (grasses) in air by time series analysis. Pollen Spores 1987;29:305–22.

Norris-Hill J. The modelling of daily Poaceae pollen concentrations. Grana 1995;34:182–8.

O'Rourke MK. Relationships between airborne pollen concentrations and weather parameters in an arid environment. Comtois P, editor. Symposium Volume, Aerobiology, Health, Environment, Montreal, June 1988:55–76.

Reiss NM, Kostic SR. Pollen season severity and meteorologic parameters in central New Jersey. J Allergy Clin Immunol 1976;53(6):609–14.

Stix E, Ferretti ML. Pollen calendars of three locations in Western Germany. Charpin J, Surinyach R, Frankland RW, editors. Atlas European des pollens allergisants, Sandoz, Paris, 1974:85–94.