A numerical model of birch pollen emission and dispersion in the atmosphere. Description of the emission module

M. Sofiev · P. Siljamo · H. Ranta · T. Linkosalo · S. Jaeger · A. Rasmussen · A. Rantio-Lehtimaki · E. Severova · J. Kukkonen

Abstract A birch pollen emission model is described and its main features are discussed. The development of the model is based on a double-threshold temperature sum model that describes the propagation of the flowering season and naturally links to the thermal time models to predict the onset and duration of flowering. For the flowering season, the emission model considers ambient humidity and precipitation rate, both of which suppress the pollen release, as well as wind speed and turbulence intensity, which promote it. These dependencies are qualitatively evaluated using the aerobiological observations. Reflecting the probabilistic character of the flowering of an individual tree in a population, the model introduces relaxation functions at the start and end of the season. The physical basis of the suggested birch pollen emission model is compared with another comprehensive emission module reported in literature. The emission model has been implemented in the SILAM dispersion modelling system, the results of which are evaluated in a companion paper.

Keywords Birch pollen · Pollen emission · Pollen forecasting · Dispersion modelling

Introduction

Pollen grains and spores have been known to disperse over long distances since the middle of the previous century (Erdtman 1931, 1935; Gregory 1961); however, it attracted greater attention only comparatively recently. The long-range transport of pollen and spores has two evident consequences: (i) short-term (hourly or daily) changes in the pollen concentrations over receptor regions, which cannot be predicted using only local observations, and (ii) a large-scale redistribution of genetic material along the atmospheric pathways (Lindgren et al. 1995). Recently, substantial progress has been made in understanding the first phenomenon (Corden et al. 2002; Damialis and Gioulekas 2005; Hjelmroos 1992; Latalova et al. 2002; Mahura et al. 2007; Ranta and Satri 2007; Rantio-Lehtimaki 1994; Siljamo et al. 2008c; Skjøth et al. 2008; Sofiev et al. 2006a). It was shown that, although the features of each specific long-range transport episode vary widely, there may be a systematic pattern in the springtime pollen redistribution in Europe: prevailing transport directions, main source regions and regions mainly receiving pollen, etc. There were several attempts to reveal...
such a pattern via a multi-annual analysis (Damialis and Gioulekas 2005; Siljamo et al. 2008a, c; 2006; Skjøth et al. 2009; 2008; 2007; Smith et al. 2008; Sofiev et al. 2006a; Stach et al. 2007), and the amount of material and the number of analysed cases are growing. However, the overall picture for Europe (i.e., systematic surveys of the main transport directions, frequency of long-range transport events, characteristic dispersion distances, source-receptor matrices, etc.) is still to be drawn.

Different taxa exhibit varying potential to the long-range transport of their pollen. Arguably, one of the most important allergenic species in Europe is birch. It is a strong allergy-provoking plant that sensitises a substantial fraction of the population in nearly all parts of Europe. In Northern Europe, it is the most abundant allergenic pollen type with approximately 15% of the population being sensitized to its allergens (WHO 2003). The distribution of silver birch (*Betula pendula* Roth.) and downy birch (*B. pubescens* L.) covers a wide area extending from the mountain regions of Southern Europe to the northernmost Fennoscandia and through Siberia to the east coast of Asia (Atkinson 1992; OECD 2003).

The pollen long-range transport has one peculiarity. It has been known for decades that the bulk of pollen is deposited near the source plant (Raynor et al. 1970; Tampieri et al. 1977; Wright 1952, 1953). However, birches as well as the other species (*Alnus, Carpinus, Corylus, Ostrya, Fagus, Quercus, Castanea*) belonging to the order Fagales are wind-pollinated trees generating vast amounts of pollen to ensure a sufficient level of fertilization of female flowers over receptor regions. Their pollen grains are quite small and light to facilitate the transportation of a substantial fraction (up to 1% for some plumes) of the released material over thousands of kilometres if weather conditions are suitable (Sofiev et al. 2006b). Since the total number of the released pollen grains is high and the threshold levels for provoking symptoms in sensitized people are quite low, the long-range transported pollen may still have a substantial health impact (Viander and Koivikko 1978).

A key tool for analysing the pollen distribution in air is the atmospheric dispersion modelling. Many studies address the local distribution of pollen and seeds (Arritt et al. 2007; Aylor et al. 2006; Kuparinne et al. 2007), mainly addressing the distribution of the genetically modified species and the possibility of unwanted spread of these plants. At a larger scale, the integrated approaches based on dynamic models that cover the main parts of the pollen life cycle and its atmospheric transport are being developed in several countries.

The European-scale operational System for Integrated modelLing of Atmospheric coMposition (SILAM) presented in the present paper has been adapted to pollen computations by an international consortium within the scope of the POLLEN project. Its pollen source term has been accepted for the allergenic computations by the MACC, a European modelling consortium. Various versions of the model have been used for forecasts of pollen distribution in Europe starting from 2005 (Sofiev et al. 2006a). Additionally, the system has been applied to the reanalysis of flowering seasons starting from 1997 (Siljamo et al. 2008c; Veriankaitė et al. 2010).

A regional-to-local scale system was started within the ASTHMA project in Southern Europe. Together with SILAM, these systems formed a demonstration service in the ESA PROMOTE project (GMES service element) from where the SILAM based service continued in the EU-PASODOBLE downstream service. A system called COSMO-ART has been developed at the University of Karlsruhe (Helbig et al. 2004; Vogel et al. 2008) and is being refined by MeteoSwiss; the objective of the development is operational forecasting for Central and South Western Europe. Integrated model development is also going on in Denmark with the regional ENVIRO-HIRLAM system (Mahura et al. 2009) and local scale OML systems. Ongoing regional activities also take place in the US (for example, Efthathiou et al. 2011).

The goals of the present paper are: (i) to present a birch pollen emission model that is implemented in SILAM; (ii) to highlight the main driving processes and the uncertainties inherent in the model; (iii) to compare the new model with another pollen emission algorithm reported in literature. A companion paper by Siljamo et al., (IJB, 2012) quantifies the performance of the SILAM system with the present pollen emission module.

Input information: dispersion and flowering models, phenological and aerobiological data

Components of the SILAM dispersion model

As shown by Sofiev et al. (2006a), the main components of virtually any comprehensive chemistry transport model can be used to describe the pollen dispersion: advection with wind, mixing due to turbulence, gravitational settling (the main mechanism of pollen dry deposition), and scavenging with precipitation. The pollen model presented in this paper is constructed as a part of the SILAM modelling system (Sofiev et al. 2008). The model dynamic core includes both Lagrangian (Sofiev et al. 2006b) and Eulerian (Galperin 2000; Sofiev 2002) advection-diffusion formulations. The removal processes are described via dry and wet deposition. The dry deposition of pollen is described via gravitational settling, which for birch pollen results in the characteristic dry deposition velocity $\sim 1.2 \text{ cm s}^{-1}$ (Sofiev et al. 2006a). Advantages and problems of the dry deposition algorithms are discussed by (Kouznetsov and Sofiev 2012). The
The temperature sum thresholds for the start and end of the season is stable from year to year (Häkkinen et al. 1998; Hänninen 1990, 1995; Linkosalo et al. 2008).

According to Linkosalo et al. (2010), the cumulative fraction $R$ of pollen released from the beginning of a year until time $t$ is piecewise linear and proportional to the temperature sum $H$ during the main flowering season.

$$R(t) = \begin{cases} 0, & H(t) < H_{fs} \\ \frac{H(t) - H_{fs}}{H_{fe} - H_{fs}}, & H_{fs} < H(t) < H_{fe} \\ 1, & H(t) > H_{fe} \end{cases} \quad (before \ the \ season)$$

$$\quad \frac{H_{fs}}{H_{fe}}, \quad (season \ goes \ on)$$

$$\quad \frac{H_{fs}}{H_{fs}}, \quad (season \ is \ over)$$

The temperature sum thresholds for the start and end of the season, $H_{fs}$ and $H_{fe}$, as well as the form of the function, $H(t)$, have to be identified from observational data.

Input data

To identify the parameters in Eq. 1, a database of more than 15,000 records of dates of phenological phases across Europe has been collected (Siljamo et al. 2008b). From that database, the date of the leaf unfolding subset (the largest in the database) was taken for the present study.

Observational data on pollen concentrations in air were obtained from the European Aeroallergen Network (EAN, https://eupolleninfo.eu/Ean/) which receives the data from around 35 countries and 300 sites all over Europe. For 2006 (the year highlighted in this paper), the set includes 5,787 daily data points from 213 stations.

The temperature time series for phenological and aerobiological stations were extracted from the meteorological archive of the European Re-Analysis (ERA-40, http://www.ecmwf.int). The ERA-40 data cover the period from 1957 until 2001 with the six-hour time resolution of the analysis. ERA-40 combined the modelling capabilities and data from the observational networks via data assimilation, thus representing the best available knowledge on the state of the atmosphere, land, and surface.

During the main season, the actual meteorological parameters are required for predicting the pollen release rate. These data were extracted from the ECMWF archive of the operational forecast. The time resolution of this dataset was 3 hours, whereas the spatial grid had a spacing of 0.25°.

Pollution emission model

Following Sofiev et al. (2006ab), the output of the pollen emission module is described as a release flux of pollen grains $E(t, i, j)$: the number of grains emitted from 1 m² of birch forest within 1 sec in a given model grid cell $(i, j)$. For dispersion computations, the final emission flux $E_{\text{meteo}}(t, i, j)$ is obtained from $E(t, i, j)$ by multiplication with the fraction of the birch forest $\phi(i, j)$ and the grid cell area $S(i, j)$:

$$E_{\text{meteo}}(t, i, j) = E(t, i, j)\phi(i, j)S(i, j)$$

The term $E(t, i, j)$ has to be decomposed using the double-threshold model of Eq. 1:

$$E(t, i, j) = \frac{dR}{dt}N_{\text{total}}p(t, i, j)F(\text{meteo})$$

Here, $N_{\text{total}}$ is the total number of pollen grains released from 1 m² of birch forest during the whole season, $p(t, i, j)$ is the probability of a tree in the given grid cell to flower at the specific time moment, and $F(\text{meteo})$ is the meteorology-dependent dynamic flux correction.

Probabilistic description of the emission flux

As shown by Siljamo et al. (2008b), the irreducible uncertainties in the season timing are large and exceed the meteorological turnover time (~3 days of persistence of the weather pattern at synoptic scale in Europe). As a result, any deterministic model of the flowering season is inaccurate; errors in, e.g., start of the season, will be comparable with or exceeding 3 days. Since the transport conditions are dictated by the meteorological situation, a shift in pollen release of more than that meteorological turnover time would lead to incorrect release parameters and to a different dispersion pattern of the emitted pollens. A solution implemented in SILAM describes the flowering in probabilistic...
terms. The approach is based on assuming a certain probability for an individual tree to flower during a specific day:

\[ p_f(H/H_{fs}, N/N_{total}) = p_{fs}(H/H_{fs})p_{fe}(N/N_{total}) \]  

(4)

Here the functions \( p_{fs} \) and \( p_{fe} \) describe the probability for a single tree to start and end the flowering, respectively, and \( N \) is the cumulative pollen amount released since the start of the season.

For a grid cell, \( p_{fs} \) and \( p_{fe} \) characterise the fraction of trees flowering at each particular time moment. These fractions are less than unity near the season start and end. They do not affect the total number of pollen grains released during the whole season \( N_{total} \) but rather extend the flowering period; it takes longer to release the prescribed \( N_{total} \) pollens if some trees are not flowering at the beginning and at the end of the season.

Start of the season

Before the season, with \( H \) approaching \( H_{fs} \), \( p_{fs} \) determines the gradual start of the pollen release from the grid cell. The specific shape of the \( p_{fs} \) function is uncertain and hardly important. The only crucial parameter is the duration of the flowering spin-up, i.e. the transition range \( \delta_H \) (in terms of the relative temperature sum) between mean daily emission intensities of, say, 5% and 95% in the grid cell (i.e., 5% and 95% of trees flowering inside the grid cell). This parameter is not measured, but can be transformed into an observable quantity as follows:

\[ \delta_H \approx \delta_t \frac{1}{H_{fs}} \frac{dH}{dt}
|_{H=H_{fs}} = \frac{(T_{fs} - T_{c-o})}{H_{fs}} \delta_t \]  

(5)

Here, \( \delta_t \) is the time lag between the flowering intensities of 5% and 95% and \( T_{fs} \) is the temperature at the first flowering day. From the point of view of phenological observations, \( \delta_t \) represents the uncertainty of the determination of the phenological phases. Such uncertainty has been quantified by Siljamo et al. (2008b).

For simplicity, \( p_{fs} \) is assumed to be piecewise linear with regard to the temperature sum:

\[ \begin{align*}
 p_{fs}(H/H_{fs}) & = p_{fs}(x) \\
 & = \begin{cases} 
 0, & x < 1 - \delta_H \\
 x + 1 - \delta_H, & 1 - \delta_H < x < 1 + \delta_H \\
 1, & x > 1 + \delta_H
\end{cases}
\]  

Here, \( x = H/H_{fs} \). According to Siljamo et al. (2008b), a \( \delta_H \) of around 20% can be used as a rough estimate.

The application of the blurring function (Eq. 6) results in (i) a gradual start of pollen release already when the temperature sum is approaching the threshold but is still below it; and (ii) all of the trees in a grid cell getting involved in the process somewhat later than the threshold is passed.

End of the season

The end of the season is determined based on the “open pocket” principle; in other words, the emission continues until \( N = N_{total} \) (rather than the end-of-season heat sum, which is not assumed in SILAM).

The total amount of pollen developed in catkins, \( N_{total} \), is a very uncertain parameter which is at present estimated semi-manually using the data from the previous year and introduced into the model as a prescribed fixed map. Some regional studies show the possibility of predicting this parameter based on meteorological data from the previous year (Rasmussen 2002); however, the approach is yet to be extended to the European scale.

Following the same approach as that of the \( p_{fs} \) function, \( p_{fe} \) reads as:

\[ p_{fe} \left( \frac{N}{N_{total}} \right) = p_{fe}(R) = \begin{cases} 
 1, & R < 1 - \delta_N \\
 1 + \frac{1}{\delta_N - R}, & 1 - \delta_N < R < 1 + \delta_N \\
 0, & R > 1 + \delta_N
\end{cases} \]  

(7)

A \( \delta_N \) estimate of around 20% is used in the current SILAM setup.

Start and propagation of the flowering season

To obtain \( dR/dt \), we used the most common temperature sum formulation in the thermal time models, i.e. an integral of daily temperature \( T \) above a cut-off level \( T_{c-o} \) starting from moment \( t_0 \):

\[ H(t) = \int_{t_0}^{t} \left( T(t) - T_{c-o} \right) * 1(T(t) - T_{c-o}) dt \]  

(8)

Here, \( I(x) \) is the Heaviside function that equals 0 for \( x < 0 \) and 1 for \( x > 0 \).

Then, the relative release rate becomes a piecewise linear function of temperature:

\[ \frac{dR(t)}{dt} = \begin{cases} 
 0, & H(t) < H_{fs} * (1 - \delta_H) \\
 \frac{dH(t)/dt}{H_{fs} - H_{fs}} \left( \frac{H(t) - T_{c-o} + (T(t) - T_{c-o})}{dH/dt} \right), & H_{fs} < H(t) \quad R < 1 \\
 1 + \frac{1}{\delta_N - R}, & R > 1 + \delta_N
\end{cases} \]  

(9)

Equation 9 requires three parameters: the starting temperature sum threshold \( H_{fs} \), the difference between the thresholds for the start and the end of pollen release \( \Delta H = H_{fe} - H_{fs} \), and the cut-off temperature \( T_{c-o} \).
and $\delta_N$ accounts for the probabilistic flowering description. A somewhat unorthodox introduction of the end of the season via $R$ rather than $H_{fe}$ allows for explicit consideration of the actual meteorology.

Corrections dependent on short-term meteorological conditions

During the main season, three meteorology dependent correction functions are applied to the dynamic release rate $E(t, i, j)$: for wind speed, relative humidity and precipitation rate.

Precipitation and humidity related corrections are derived from known “prohibiting” thresholds totally suppressing the pollen release. Until these thresholds are reached, these variables do not affect the release (neither do they promote it). Near the threshold, the piecewise linearly decreasing transition function is as follows:

$$f_{thr}(x, x_{low}, x_{high}) = \begin{cases} 1, & x \leq x_{low} \\ \frac{x_{high} - x}{x_{high} - x_{low}}, & x_{low} < x < x_{high} \\ 0, & x \geq x_{high} \end{cases}$$

(10)

The lower and upper thresholds of relative humidity are taken as $q_{low}=50\%$ and $q_{high}=80\%$.

For precipitation, $P_{low}=0$. For the selection of $P_{high}$, several considerations have to be taken into account. Any noticeable rain suppresses the release and scavenges out the emitted grains. The pollen release can also be stopped by high relative humidity associated with rain, which covers wider areas than the rain event itself. However, short-term convective rains cover an area much smaller than the grid cell (Morel and Senesi 2002). Such scattered precipitation still allows the trees to emit pollen from the dry parts of the grid cell area. Taking into account the above uncertainties, the estimate $P_{high}=0.5$ mm hr$^{-1}$ (the grid cell average rate) is taken as the threshold suppressing the pollen emission.

For the wind dependent correction, three phenomena need to be taken into account: (i) in the case of low wind but developed thermal convection, turbulence alone is sufficient to kick-start the release by generating sub-grid convective winds, (ii) stronger wind promotes the release by picking the pollen grains from open catkins; (iii) after reaching some level, further increase in wind strength does not affect the release rate which is then limited by the availability of ready-to-fly pollen grains in the catkins. These phenomena can be included in a single function as follows:

$$f_{wind} = f_{stagnant} + f_{promote} \left[ 1 - \exp \left( - \frac{U + w^*}{U_{satur}} \right) \right]$$

(11)

Here, $U$ is the wind speed, $w^*$ is convective velocity scale, $U_{satur}$ is the saturation wind speed, and $(f_{stagnant} + f_{promote})$ is the maximum “promotion” that wind can give to the release rate. In stagnant conditions, Eq. 11 suppresses the release by the $f_{stagnant}$ factor.

In the current SILAM version, $U_{satur}=5$ m sec$^{-1}$, $f_{stagnant}=0.5$ and $f_{promote}=1$, which implies no impact at wind speed $\sim$1 ms$^{-1}$ and three-fold stronger release at very strong wind in comparison with calm conditions.

The resulting emission rate is a product of the above described specific terms:

$$E_{model} = \begin{cases} 0 : & \begin{cases} H < H_{fe} \ast (1 - \delta_H) \\ S(i, j) \phi(i, j) N_{total} \frac{T - T_{c-o}}{\Delta H} f_{wind}(U, w^*) p_{fs}(H) p_{fe}(R) \ast f_{thr}(q, q_{low}, q_{high}) f_{thr}(P, P_{low}, P_{high}) \ast \end{cases} \\ 0 : & \begin{cases} H > H_{fe}, R < 1 + \delta_N \\ R > 1 + \delta_N \end{cases} \end{cases}$$

(12)

One of the results of accounting for these corrections is that the heat sum at the end of the season is neither predictable nor important: it is not only temperature that determines the release during the main season. The $H_{fe}$ is needed only to determine $\Delta H$.

Determination of parameters of the temperature sum model

The parameters needed for Eq. 12, namely, $H_{fe}$, $\Delta H$, $T_{c-o}$, $t_0$, were identified by fitting optimally the flowering start and end dates into the phenological and aerobiological observations. To overcome the problems related to the regional variations of these parameters, the European continent was split into 33 sub-regions which together cover its entire territory and have a limited but noticeable overlap with each other ($\sim$10% of their areas). The independent fittings were performed inside each region. The overlap between the regions resulted in partly overlapping sets of observations used by the fitting procedures in the neighbouring regions. This smoothed out the contrasts between the parameterizations in the neighbouring regions.
Since none of the above parameters is observed directly, the fitting variables were the dates of the phenological phases, such as the leaf unfolding date. These dates are not observed directly. The fitting then minimised the difference between the observed dates and the corresponding model predictions by varying the above parameters. To compute the temperature sum, we used the discrete version of the definition (Eq. 8):

$$H(D) = \sum_{d=0}^{D} (T(d) - T_{c-o}) \ast 1(T(d) - T_{c-o}) \quad (13)$$

Here, $D$ is day and the bar denotes the daily averaging constructed from the 6-hour ERA values, and $D_s$ is the starting day of the $H$ integration.

The modelled starting date of the flowering, $D_{fs}^{mod}(s)$, for the specific station $s$ was then defined as the first day when $H(D_s,s) \geq H_f$. The criterion for the fitting was the RMS of the model predictions $D_{fs}^{mod}(s)$ versus observed $D_{fs}^{obs}(s)$:

$$J_{fs}(H_{fs},D_s,T_{c-o},r) = \frac{1}{N_r} \sum_{s=1}^{N_r} (D_{fs}^{mod}(s) - D_{fs}^{obs}(s))^2$$

$$\rightarrow \min_{H_{fs},D_s,T_{c-o}} \quad (14)$$

Here, $N_r$ is the number of stations in the sub-region $r$ and $J_{fs}$ is the sub-regional cost function.

Fitting of $H_{fe}$ proceeded similarly, except for the observational dataset. Since the phenological observations of the end of flowering are scarce, we used the aerobiological data from EAN. The flowering ending date for each grid cell was computed using the 97.5% criterion for the season end (Goldberg et al. 1988), after which the heat sum was computed until this date using the same start day and temperature cut-off as determined for $H_{fs}$. Upon completion of fitting procedures for $H_{fs}$ and for $H_{fe}$, their difference, $\Delta H$, was calculated for each sub-region.

Expectedly, the optimal fitting problem was ill-posed and the uniqueness of the solution was not guaranteed, which was a roadblock because the level of noise in the observational data was sometimes overwhelming (Siljamo et al. 2008b). Some improvements were achieved owing to the interdependence of the above parameters. As it was noticed by Linkosalo et al. (2000), $D_s$ and $T_{c-o}$ could be assumed, so that the fitting has to be done only for the temperature sum thresholds. The resulting value of the cost function $J_{fs}$ was practically the same as for the fitting with all three parameters varying.

The assumed values for birch were $T_{c-o} = 3.5^\circ C$ and $D_s = 60$ (March 1). There is certainly an ambiguity in these values; one can argue, for example, that March 1 is too early in the north and may be already quite late in the south, i.e. there has to be latitude dependency of at least $D_s$. In the present parameterization, however, all such trends are automatically reflected in the heat sum threshold map. Introducing any further complexity to the fitting algorithm has shown to bring no gain in terms of the final cost function value. As an additional indication of March being a good choice for pan-European starting date, Linkosalo and Lechowicz (2006) found a weak but noticeable effect of the light conditions triggering the birch leaf bud development. In such cases, the dates near spring equinox are particularly good for starting the integration because the daylight time is the same for the whole Europe.

The quality of the fitting outcome can be evaluated using several criteria. First, the residual of the fitting should be larger than the objective uncertainty of the observations: $J > \sigma^2_{obs}$. This requirement ensures that the model is not over-fitted to noise in the data. Second, the residual should be smaller than the sum of the observational uncertainty and the inter-annual variability of the flowering dates: $J < \sigma^2_{obs} + \sigma^2_{annual}$. This means that the model resolves this variability. Third, the large-scale features of the observed and fitted spatial patterns of the flowering dates should be similar, but the high-frequency fluctuations in the data map should be smoothed out by the model. Finally, the threshold map can be analysed using the classical work by Linsser (1867). According to Linsser’s law, for any heat-driven phenological phenomenon, a heat-sum threshold is a constant fraction of the overall accumulated effective temperature sum (ETS) over the whole growing period. In other words, the map of the heat sum threshold should be proportional to the map of ETS (the latter one can be easily computed from meteorological data over the period of active vegetation).

The results of the fitting (Fig. 1) satisfy all above requirements. The pattern is smooth with gradual decrease towards the north, i.e. the scatter of the input data is efficiently smoothed out. Figure 2 shows that the error of the start of flowering is indeed comparable but larger than the irreducible uncertainty in the phenological data itself, as estimated by Siljamo et al. (2008b) (these, in particular, include both natural variability of the flowering timing for different trees and the representativeness error for phenological observations). Accordingly, the procedure is conservative enough to avoid over-fitting the model to the noise in the data.

The fitting error (Fig. 2a) is lower than the inter-annual variability of the flowering dates which for some years and regions can span several weeks. For instance, in the region of the Gulf of Finland the flowering in 1999 was about 12 days earlier than the average over 1970–2003, in comparison with 6 days of the fitting RMS. Therefore, we conclude that the suggested fitting utilises (most) useful information stored in the phenological records.
The difference between the temperature sum threshold for the start of the season $H_{fs}$ is more than a factor of three between Southern and Northern Europe. The spatial trends of the observational uncertainty and residuals (Fig. 2) are similar, which confirms that the main contributor to the residuals is the observational uncertainty itself. Finally, the ratio of $H_{fs}$ to ETS computed for 2006 (Fig. 3) is indeed nearly constant over the European continent. Exceptions include the mountain areas and Northern Lapland, for which the phenological information was almost non-existent and the accuracy of both Lensser’s law and our fitting is questionable. Therefore, all above mentioned quality criteria are satisfied.

Fitting of $H_{fe}$ was less straightforward because the end of the season had to be defined from aerobiological observations which are sensitive to the long-range transport of pollen. Since observations cannot distinguish between the LRT pollen grains and those produced locally, the flowering...
duration tends to be overestimated. In addition, the temperature sum is accumulated throughout the whole season, including rainy and high-humidity days during which the release does not take place, which is another uncertainty pushing towards overestimation. Therefore, the $H_f$ map in Fig. 1 will require adjustments (see discussion in the companion paper). The empirical value $\Delta H=50$ degree-days has been shown to provide acceptable results.

There are some peculiarities in the threshold maps. Thus, both the start and end of season thresholds are at their highest in marine climate which can be quite warm in early spring but considerably colder during the main flowering season in April–May. An explanation of this behaviour can be the suppressed diurnal variation of temperature in the coastal regions and the much slower warming-up of the environment, which makes the cut-off temperature uncertain and probably different from the more continental conditions. As a result, the single-parameter Europe-wide fitting may not be appropriate for those regions: the cut-off temperature should possibly be taken somewhat higher than in a continental climate.

The other peculiarity is the very low thresholds in Southern (for $H_f$) and Central (for $H_e$) Norway. Despite the fact that these are coastal regions, the values there are lower than even in the Finnish Lapland. This suspicious behaviour can possibly originate from the large altitude variability in the mountainous regions and the sharp rise in the topography from the sea level up to the ridges. The ERA-40 temperature field and the fitting procedure do not resolve the narrow
valleys and the quick rise of the relief. The start of the flowering of trees in the valleys is then driven by local temperatures which may have little in common with the mean temperature averaged over the 1.125° × 1.125° ERA grid cell. Also, there are too few reported observations for Norway. A similar effect shows up in the Alpine regions where the valley temperatures and the corresponding thresholds are evidently higher than up in the mountains. These regions are also outliers in the Lenssers’s ratio map in Fig. 3. To resolve these features, the grid cell should be a couple of kilometres in size, which lies outside the scope of the European model development study.

**Evaluation of the process representations in the emission module**

The main season timing and propagation model

There exist several approaches for predicting the dates of phenological phases. Our selection of the thermal type model is based on findings of Linkosalo et al. (2008), who compared the thermal time, sequential (Hänninen 1990), parallel (Cannell and Smith 1983) and flexible models (Chuine 2000; Schaber and Badeck 2003) of leaf bud burst against an independent control dataset and found that the simplest thermal time models performed better than the more complicated approaches. They suggested that the complex phenological models were over-parameterized and able to adapt to noise in the learning dataset in addition to the phenological phenomenon itself. Therefore, the simple form of Eq. 1 was considered optimal.

The second challenge was the determination of the threshold values in the Eq. 1. There are numerous parameterizations for $H_f$ (Cannell and Smith 1983; Chuine 2000; Hänninen 1990; Linkosalo et al. 2008; Menzel 1997; Rotzer and Chmielewski 2001; Schaber and Badeck 2003). However, they appeared practically incomparable with each other and applicable only to the regions and species for which they were developed; both the temperature sum threshold and the formulations for $H(t)$ vary from model to model, even if their regions overlap. Attempts to generalise the formulations by introducing, for example, the latitudinal dependence of $H_f$ did not resolve the problem either. Thus, for instance, the temperature sum threshold formula with the explicit latitudinal dependence for Germany (Menzel 1997) cannot be extrapolated to the Finnish latitudes, as it would lead to negative threshold values. As a result, none of the existing thermal time model parameterizations was found directly applicable to the European-wide applications.

With no suitable model applicable for the whole of Europe, the parameters of Eq. 1 have been identified afresh, as described in the previous section. As a by-product, the identification procedure explained the differences between the existing local parameterizations; with the model parameters being interdependent, one can choose up to two of them offsetting the introduced errors with optimal selection of the remaining one(s). Since the range of reasonable
choices of the assumed parameters is wide, comparison of the obtained parameterizations can be problematic. For the present model, we just unified these assumed values for $T_{c,o}$ and $D_c$.

Short-term variations of pollen release

In this section, we analyse the short-term response of the emission module to meteorological forcing via comparison

Fig. 4 Dependence of the daily observed (left), daily predicted (middle) and hourly predicted (right) counts [m$^{-3}$] of birch pollen on the meteorological forcing: rain intensity (first row, mm hr$^{-1}$), relative humidity (second row, %), temperature at 2 m above the surface (third row, °C), and wind speed at 10 m above the surface (fourth row, m s$^{-1}$). Colours show the fraction [%] of values falling into the corresponding range of values.
with observations. The emission fluxes are not measured explicitly, therefore the analysis has to be based on the pollen counts during the main pollen season compared with the SILAM predicted concentrations. For this evaluation, we took the EAN data for 2006 and ran SILAM with a 30 km resolution over Europe using the ECMWF meteorological data.

Such evaluation is evidently semi-qualitative and its numerical outcome should be taken with care. Particularly uncertain parts refer to temperature and relative humidity; daily observations miss the strong diurnal cycles of these parameters. The same also refers to wind, although to a somewhat lesser extent.

In Fig. 4, the four main meteorological parameters extracted from the meteorological model fields are plotted against the observed and predicted pollen counts at aerobiological sites. As seen from the upper row, rain suppresses the pollen release, which results in a sharp decline of the predicted concentrations, in agreement with the observations. The substantial concentrations predicted and observed in a small fraction of cases should originate from the regional transport (in the model, the local emission is fully stopped for \( P > 0.5 \) mm hr\(^{-1}\)).

Dependence on relative humidity is more complicated (Fig. 4, second row). Since the daily concentrations are plotted against humidity picked at 12:00 UTC, there are few cases with RH >80%, and the counts are indeed low during these days. A tendency of lower counts for higher RH is visible in both daily plots but the scatter is high. The diurnal variation of RH is resolved in the hourly scatter plot at the right-hand side, which reveals that the high-humidity hours (RH >90%) are characterised by very low pollen counts; the intermediate-humidity levels of 90% > RH > 60% correspond to intermediate concentrations, whereas in most cases dry periods correspond to high counts.

The correlation of the pollen counts with temperature (Fig. 4, third row) is moderate. However, it is well seen that for the hours with high temperatures the model tends to predict higher pollen concentrations, whereas very low temperatures correspond to a low pollen load—both in predicted and observed plots.

The weakest correlation is seen for wind speed; the overall scattering is very large in both observed and predicted fields, which is the result of two competitive effects. The stronger wind speed increases the emission rate but also improves ventilation and promotes turbulent mixing. The net effect is then practically negligible, as seen from both modelled and observed plots.

The scatter plots in Fig. 4 also illustrate the mechanisms responsible for the diurnal variation of emission. Indeed, formulations (Eq. 12) do not include forcing dependence on time of the day. Birch trees do not have any sun- or light-following mechanism that would control emission of pollen (contrary to, e.g., sunflowers). Therefore, the high humidity, low temperature and low wind speed are the only factors reducing the emissions during night and causing the diurnal variation of the rates. A commonly accepted zero level of emission during the night-time can thus be reached in the model only if humidity is above the upper threshold or temperature is below the cut-off limit. In Fig. 4, such cases, whether occurring during day or night, were characterised by quite low pollen concentrations, which could still reach dozens of pollen grains per m\(^3\), which is in agreement with the observations.

**Comparison with another emission parameterization**

In this section, we compare the formulations of the SILAM pollen emission module with the birch pollen emission parameterization developed by Helbig et al. (2004) further referred to as H04. Apart from the present model, the H04 approach seems to be the only comprehensive parameterization of pollen emission applicable at a regional-to-continental scale. It was implemented in the COSMO-ART modelling system and used for, e.g., modelling of the birch pollen episode in Switzerland in 2006 (Vogel et al. 2008). Recently, a combination of H04 with the SILAM formulations has been applied in the United States to birch and ragweed simulations (Efstathiou et al. 2011).

The principal difference between the approaches is that the H04 algorithm computes the emission flux as a product of a characteristic velocity scale with a characteristic pollen concentration, adjusted with correction functions dependent on other parameters. The flux is then driven by the turbulent stress, whereas our approach follows the temperature-driven model by Linkosalo et al. (2010). As a result, the H04 approach has to utilise the assumed duration of the flowering (included in the formulation of the probability of trees to bloom), while the current model follows the actual meteorology-driven developments.

The characteristic velocity scale in H04 is taken as friction velocity \( u^* \), which is useful as a measure of the mechanically induced turbulence near the surface. However, it may not be the ideal parameter describing the mechanical stress to the birch flowers, which are several meters from the surface and located in the tree crown where the similarity theory is not applicable. At this height, the stress could be represented by regular wind blow through the canopy rather than turbulence. The closest standard meteorological variable would be the wind speed at 10 metres above the displacement height (in other words, near the tree top), which is the variable used in SILAM. It is a diagnostic variable in all meteorological models and incorporates indirectly the horizontal wind shear stress, thus making the involvement of \( u^* \) unnecessary.

The turbulence-driven stress would become dominant in free convection conditions, in other words, in unstable
stratification and low mean wind conditions. The corresponding term is absent from H04, whereas the present approach includes it via the convective velocity scale $w^*$. The impacts of precipitation and relative humidity to the emission rate are similar in both parameterizations. Temperature is also an emission promoter in H04 formulations; however, it does not play as central a role as in our model.

One of the key factors in H04 is the leaf area index (LAI), an increase which reduces the flowering intensity. We did not include this parameter at all because its relation to the birch flowering stage is neither straightforward nor easy to determine. The usual presumption that the flowering finishes before the leaves reach full size (in order not to inhibit the pollen dispersal) corroborates the H04 approach. Nevertheless, in the dataset (Siljamo et al. 2008b) the bud burst day was often before the first flowering day. As a result, LAI might start rising already before the flowering begins. However, LAI can still be useful in the case of prolonged flowering in wet, cold weather when leaves grow up before all pollen grains are released.

Another process not included in the current parameterization but considered in H04 is pollen resuspension. As noted by H04, this process is very poorly known and can only take place when the wind is very strong and gusty (above 15 m sec$^{-1}$). Its influence can possibly be noticed at the end of the season, in flat terrains and in dry conditions only. Therefore, we considered it as a rare phenomenon, the uncertainty of which is larger than the potential gain of its inclusion into the model.

Summary

The suggested pollen emission model for birch follows the concept of thermal time phenological models and, in particular, the double-threshold temperature sum approach determining the development of flowering during the whole spring season.

Apart from temperature, pollen release rate is modulated by ambient humidity, precipitation and wind speed. Higher humidity and rain suppress the release, whereas stronger wind promotes it. Atmospheric turbulence is taken into account via the turbulent velocity scale and thus becomes important only in cases of near-free convection.

The probability of an individual tree to enter the flowering stage is considered via the uncertainty of the temperature sum threshold determining the start of flowering.

The end of the season is described via the open pocket principle, according to which flowering continues until the initially available amount of pollen is released.

Numerical values of the model parameters—temperature sum threshold for the start and end of flowering, critical levels of relative humidity and precipitation intensity, and characteristic wind speed—were identified via optimal fitting to the European phenological and aerobiological data.

The model does not include the explicit diurnal variation forcing, which is obtained as a by-product of the meteorological forcing. Pollen resuspension and the relation between flowering and leaf area index are also left out of the parameterization due to high uncertainties in both parameters and their unclear relation to the pollen release processes.

The model processes have been qualitatively evaluated using the observed pollen counts in 2006, which were related to the meteorological input data and compared with the model predictions. The quantitative analysis of the model performance is the subject of the companion paper.

The described model is freely available from the SILAM model website at http://silam.fmi.fi.

Acknowledgments

This work was performed within the scope of the POLLEN project of the Academy of Finland and supported by the FP7 projects HIALINE, MACC and PASODOBLE. Networking activities were supported by the COST-ES0603 EUPOL Action. The pollen observational data of the European Aeroallergen Network (EAN) are kindly acknowledged.

Open Access

This article is distributed under the terms of the Creative Commons Attribution License which permits any use, distribution, and reproduction in any medium, provided the original author(s) and the source are credited.

References

Arritt RW, Clark CA, Goggi AS, Sanchez HL, Westgate ME, Riese JM (2007) Lagrangian numerical simulations of canopy air flow effects on maize pollen dispersal. Field Crops Research 102:151–162. doi:10.1016/j.fcr.2007.03.008

Atkinson MD (1992) Betula pendula Roth (B. verrucosa Ehrh.) and B. pubescens Ehrh. J Ecol 80:837–870

Aylor DE, Boehm MT, Shields EJ (2006) Quantifying aerial concentrations of maise pollen in the atmospheric surface layer using remote-piloted airplanes and lagrangian stochastic modelling. J Applied Meteorol Climatol 45:1003–1015

Cannell MGR, Smith RI (1983) Thermal time, chill days and prediction of budburst in Picea sitchensis. J Appl Ecol 20:951–963

Chuine I (2000) A unified model for budburst of trees. J Theor Biol 207:337–347

Corden JM, Stach A, Millington W (2002) A comparison of Betula pollen season at two European sites; Derby, United Kingdom and Poznan, Poland (1995–1999). Aerobiologia 18:53–54

Damialis A, Gioulekas D (2005) Transport of airborne pollen into the city of Thessaloniki: the effects of wind direction, speed and persistence. Int J Biometeorol 49:139–145. doi:10.1007/s00484-004-0229-2

Efstathiou C, Isukapalli S, Georgopoulos P (2011) A mechanistic modeling system for estimating large-scale emissions and transport of pollen and co-allergens. Atmospheric Environment 45 (13):2260–2276. doi:10.1016/j.atmosenv.2010.12.008. Available at: http://dx.doi.org/10.1016/j.atmosenv.2010.12.008

Erdman G (1931) Pollen-statistics: A new research method in paleo-ecology. Science 73:399–401
Skjøth CA, Sommer J, Brandt J, Hvidberg M, Geels C, Hansen KM, Hertel O, Frohn LM, Christensen JH (2008) Copenhagen—a significant source of birch (Betula) pollen? Int J Biomeeteorol 52:453–462. doi:10.1007/s00484-007-0139-y

Skjøth CA, Smith M, Brandt J, Emberlin J (2009) Are the birch trees in Southern England a source of Betula pollen for North London? Int J Biomeeteorol 75–86. doi:10.1007/s00484-008-0192-1

Smith FB, Clark MJ (1989) The transport and deposition of radioactive debris from the Chernobyl nuclear power plant accident with special emphasis on consequences to the United Kingdom. Meteorological Office Scientific Paper 42

Smith M, Skjøth CA, Myszkowska D et al (2008) Long-range transport of Ambrosia pollen to Poland. Environ Res 148:1402–1411. doi:10.1016/j.jagreformet.2008.04.005

Sofiev M (2002) Extended resistance analogy for construction of the vertical diffusion scheme for dispersion models. J Geophys Res–Atmos 107(D12). doi:10.1029/2001JD001233

Sofiev M, Siljamo P, Ranta H, Rantio-Lehtimaki A (2006a) Towards numerical forecasting of long-range air transport of birch pollen: theoretical considerations and a feasibility study. Int J Biometeorol 50(6):392–402. doi:10.1007/s00484-006-0027-x

Sofiev M, Siljamo P, Valkama I, Ilvonen M, Kukkonen J (2006b) A dispersion modelling system SILAM and its evaluation against ETEX data. Atmospheric Environ 40(4):674–685. doi:10.1016/j.atmosenv.2005.09.069

Sofiev M, Galperin MV, Genikhovich E (2008) Construction and evaluation of Eulerian dynamic core for the air quality and emergency modeling system SILAM. In: Borrego C, Miranda AI (eds) NATO Science for peace and security Series C: Environmental Security. Air pollution modelling and its application, XIX. Springer-Verlag, Berlin, pp 699–701

Stach A, Smith M, Skjøth CA, Brandt J (2007) Examining Ambrosia pollen episodes at Poznań (Poland) using back-trajectory analysis. Int J Biometeorol 275–286. doi:10.1007/s00484-006-0068-1

Tampieri F, Mandrioli P, Puppi GL (1977) Medium range transport of airborne pollen. Agric Meteorol 18:9–20

Veriankaitė L, Siljamo P, Sofiev M, Sauliene I, Kukkonen J (2010) Modelling analysis of source regions of long-range transported birch pollen that influences allergenic seasons in Lithuania. Aerobiologia 26(1):47–62. doi:10.1007/s10453-009-9142-6

Viander M, Koivikko A (1978) The seasonal symptoms of hyposensitized and untreated hay fever patients in relation to birch pollen counts: correlation with nasal sensitivity, prick tests and RAST. Clinical Allergy 8:387–396

Vogel H, Pauling A, Vogel B (2008) Numerical simulation of birch pollen dispersion with an operational weather forecast system. Int J Biometeorol 52(8):805–814. doi:10.1007/s00484-008-0174-3

WHO (2003) Phenology and human health: allergic disorders. WHO, Copenhagen

Wright JW (1952) Pollen dispersion of some forest trees. U.S. Department of Agriculture, Forest Service, Northeastern Forest Experiment Station, Upper Darby, PA

Wright JW (1953) Pollen dispersion studies: Some practical applications. J Forestry 114–118