Geoscientific
Model Development

EMPOL 1.0: a new parameterization of pollen emission in numerical weather prediction models

K. Zink[1], **A. Pauling**[1], **M. W. Rotach**[2], **H. Vogel**[3], **P. Kaufmann**[1], **and B. Clot**[1]

[1]Federal Office of Meteorology and Climatology MeteoSwiss, Zürich, Switzerland
[2]Institute for Meteorology and Geophysics, University of Innsbruck, Innsbruck, Austria
[3]Institute for Meteorology and Climate Research, Karlsruhe Institute of Technology, Karlsruhe, Germany

Correspondence to: K. Zink (katrin.zink@meteoswiss.ch) and A. Pauling (andreas.pauling@meteoswiss.ch)

Abstract. Simulating pollen concentrations with numerical weather prediction (NWP) systems requires a parameterization for pollen emission. We have developed a parameterization that is adaptable for different plant species. Both biological and physical processes of pollen emission are taken into account by parameterizing emission as a two-step process: (1) the release of the pollen from the flowers, and (2) their entrainment into the atmosphere. Key factors influencing emission are temperature, relative humidity, the turbulent kinetic energy and precipitation.

We have simulated the birch pollen season of 2012 using the NWP system COSMO-ART (**Co**nsortium for **S**mall-scale **Mo**delling – **A**erosols and **R**eactive **T**race Gases), both with a parameterization already present in the model and with our new parameterization EMPOL. The statistical results show that the performance of the model can be enhanced by using EMPOL.

1 Introduction

A relatively high proportion of the population in industrialized countries suffer from pollen allergies. Different allergenic plant species shed their pollen at different times of the year, leading to several pollen seasons (e.g., birch pollen season, grass pollen season). In central Europe, the most important periods for patients are the tree pollen season in spring, the grass pollen season in late spring and early summer, and the ragweed pollen season in late summer and autumn (only in regions where this invasive plant is present). Even though medication is possible, the best way to reduce allergic

symptoms remains a complete avoidance of the allergens (van Moerbeke, 1997). Therefore, it is of great importance to forecast the distribution of airborne pollen a few days in advance. Currently, pollen forecasts are mainly based on pollen monitoring, weather forecasts, climatological information about the pollen seasons, and the experience of the forecaster. The available technology (Hirst type traps, Hirst, 1952) for pollen monitoring demands a great deal of manpower since the pollen grains have to be counted manually. Therefore, the density of the pollen sites is low compared to that of meteorological sites. Thus, the spatial resolution of pollen forecasts is very low (essentially, forecasts are only available at the observational sites themselves).

Recently, pollen dispersion has been integrated into numerical weather prediction (NWP) models. The knowledge of the simulated spatial and temporal evolution of pollen concentrations enables the forecaster to make nationwide predictions instead of forecasts for the pollen sites only. The prerequisite for reliable numerical simulations of pollen concentrations are (i) an up-to-date distribution map of the plant that indicates the pollen sources, (ii) a NWP system that can deal with particles (including transport, deposition, washout of the particles etc.), and (iii) a parameterization of the emission process.

Assuming that the first two prerequisites – the distribution map of the plant and the NWP system including particle dispersion – are available, this publication is focused on the emission of pollen grains, which is a combination of physical and biological processes that are characteristic for each plant species. Thus, the emission parameterization has to address two questions: (i) When does emission take place? (ii) How

many pollen grains are emitted? To answer these questions, we have to look more closely at the biological and physical processes leading to pollen ripening and release. Two time scales can be distinguished that require the use of two sub-models within the emission parameterization:

1. The seasonal cycle of pollen emission depends on the number of mature flowers or inflorescences, and hence on the development of the plants. It is driven by the weather conditions during the preceding weeks or months and usually is described via phenological models (e.g. Sarvas, 1974; García-Mozo et al., 2009).

2. The diurnal cycle of pollen emission is driven by the current meteorological conditions that lead to a rupture of the anthers (thus, the release of pollen grains from the flowers) and to the entrainment of the pollen grains into the atmosphere. This process happens in a time frame of seconds to hours.

Although the processes leading to pollen emission have been described in the literature, the information that can be found is mainly of a qualitative nature (e.g. Bianchi et al., 1959). Additionally, both the development of the plants and the release of the pollen are processes that are specific to each plant species. Hence, an emission parameterization that is supposed to work for different plant species has to be very flexible regarding these biological and physical processes.

In Sect. 2, we briefly describe the meteorological model that is used in our study. A comparison between different pollen emission parameterizations is given in Sect. 3. In Sect. 4, a new parameterization for pollen emission is developed, followed by a description of a basic tuning. The application of this parameterization and a comparison of the results with an existing parameterization can be found in Sects. 5 and 6. Finally, Sect. 7 contains the summary and conclusions.

Throughout the publication, we have used the following terms to describe the different processes and parameters in the context of pollen emission:

– *anthesis:* opening of the anthers.

– *pollen release:* release of the pollen from the anthers, either into a reservoir or directly into the atmosphere.

– *pollen presentation:* making the pollen grains available for entrainment into the atmosphere. Usually, this is the result of the combined processes of anthesis and pollen release.

– *entrainment:* uplifting of the pollen from the reservoir or from the anthers into the atmosphere.

– *pollen emission:* the combined processes of pollen release and entrainment into the atmosphere. This term is used when the different steps during the emission

process are not distinguished. This formulation is necessary since many emission parameterizations treat pollen release and entrainment as a single process.

– *pollen production:* the amount of pollen that is produced per square meter and per year.

– *pollen concentration:* the concentration of airborne pollen, given in number of pollen grains per cubic meter of air.

2   The NWP model system COSMO-ART

COSMO is a non-hydrostatic regional NWP model that is used for operational weather forecasts in various European countries (Steppeler et al., 2002). Vogel et al. (2009) have developed an extension ART (**A**erosols and **R**eactive **T**race Gases) to COSMO in order to study the interaction between aerosols and the atmosphere. Information about the code availability can be found in the Supplement.

Physical processes that are incorporated into COSMO-ART include transport by the mean wind, turbulent diffusion, dry and wet deposition, coagulation, condensation, washout, and sedimentation of the aerosols and reactive trace gases. ART includes, amongst others, a module to simulate the emission and dispersion of pollen grains (Vogel et al., 2008; Zink et al., 2012). The default parameterization of pollen emission follows the suggestions of Helbig et al. (2004) and Vogel et al. (2008).

At MeteoSwiss, COSMO-ART has been used as an operational tool for birch pollen forecasting since 2011, using a modified version of the Helbig et al. (2004) emission parameterization. The modification includes the description of the pollen season as well as the influences of temperature, humidity, and wind speed on the pollen emission. It is designed to better model the physiological processes in the plants. Since the quantitative relationships between the meteorological parameters and pollen emission are largely unknown, the implementation of the qualitative knowledge about plant physiology is somewhat subjective. The error between modeled and observed pollen concentrations was evaluated as a function of temperature, relative humidity and wind speed. Using these results, the meteorological functions in the emission parameterization of Helbig et al. (2004) were adapted. This was done based on observational data from Switzerland.

3   Available emission parameterizations

The approaches to parameterize pollen emission that can be found in the literature differ greatly in complexity. Very simple solutions use spatially and temporally uniform emission fluxes (e.g. Pasken and Pietrowicz, 2005). More sophisticated versions include current meteorological conditions and/or a curve representing the typical pollinating season

(e.g. Helbig et al., 2004; Schueler and Schlünzen, 2006; Marceau et al., 2011; Sofiev et al., 2013). In addition to the current meteorological conditions, the model of Martin et al. (2010) includes previous values of relative humidity.

In this section, we describe the differences and similarities of three different emission parameterizations in some detail. These parameterizations are:

- The parameterization of Helbig et al. (2004), in the following referred to as "H_{orig}". It was used to simulate both birch (Vogel et al., 2008) and ragweed (Zink et al., 2012) pollen concentrations.

- The parameterization that has been used for operational numerical birch pollen forecasts at the Federal Office of Meteorology and Climatology in Switzerland (MeteoSwiss), an optimized formulation of H_{orig}. In the following referred to as "H_{opt}" (compare Sect. 2).

- The parameterization of Sofiev et al. (2013). In the following referred to as "S13". It has been implemented into the System for Integrated modeLling of Atmospheric coMposition (SILAM) and was applied for birch pollen emission.

These are the only comprehensive parameterizations of which we are aware that are incorporated into NWP systems for the application to wider regions. The comparison is divided into three parts reflecting the nature of the parameterizations: the first part describes the seasonal cycle, the second part the daily cycle of pollen emission. Part three explains additional features of the parameterizations. A table summarizing the differences between the parameterizations can be found in the Supplement.

3.1 Description of the pollen season

The amount of pollen that is ripe and available for emission depends greatly upon the state of the pollen season. At the beginning and the end of the season only a few plants flower, hence, the amount of available pollen grains is small, regardless of the meteorological conditions. It is therefore essential to introduce a mathematical model that describes the course of the pollen season on a daily basis. H_{orig} assumes the birch pollen season to last for 30 days and to have a fixed shape of a parabola. In H_{opt}, a more sophisticated description of the pollen season is used. Based on measured pollen data gathered in Switzerland, the shape of the pollen curve is chosen to be positively skewed. A similarly skewed shape of the pollen curve has been reported for central Europe by Grewling et al. (2012). Additionally, the length of the season is variable depending on the temperature during the pollen season. This takes into account the fact that a warm spring season results in a short and intense birch pollen season, whereas a cool spring tends to lengthen the birch pollen season because not all birch trees flower at the same time. S13 introduces an internal model that reflects the seasonal pollen curve. It uses

temperature sums to predict the start and course of the pollen season. The season ends when a certain amount of pollen has been emitted. The upswing and downswing of the pollen curve is parameterized using relaxation functions describing the probability for a single tree to bloom at a certain day.

3.2 Meteorological influences

Naturally, emission has to be linked to a velocity scale since the pollen grains have to be lifted into the air by wind currents. In H_{orig} and H_{opt}, the friction velocity u_* is taken as the parameter influencing the amount of emitted pollen. A threshold friction velocity u_{*t} has to be reached in order to allow emission. The value of u_{*t} is derived using a parameterization for dust entrainment and a meteorological correction factor (see below). Emission by free convection is not taken into account. S13 uses both the 10 m wind speed and the convective velocity scale w_* to take into account both ways of pollen entrainment into the atmosphere.

Theoretically, if unlimited amounts of pollen were available, higher wind speeds would yield stronger entrainment, and hence more airborne pollen. In reality, this is limited by the fact that at a certain point, the flowers will run out of pollen grains. This is not considered in H_{orig}. H_{opt} and S13 take this into account by introducing a threshold wind speed or a function converging to a maximal value that stops further increase of emission.

Short-term pollen emission is driven by current meteorological conditions. All three parameterizations consider these effects using correction terms. H_{orig} and H_{opt} use meteorological correction terms for temperature, relative humidity and wind speed that influence the value of the threshold friction velocity u_{*t}. However, the individual terms are different for both parameterizations (cutoffs at different thresholds etc.). Both H_{orig} and H_{opt}, consider precipitation through washout. S13 considers relative humidity and precipitation as hindrances for emission if their values are in a certain range. In S13, temperature plays an important role through the seasonal description, short-term effects on emission are not taken into account.

3.3 Other features

H_{orig} and H_{opt} use plant-dependent parameters such as the leaf area index (LAI), and the height of the plants that influence the amount of emitted pollen. Their use in the model is reciprocal, leading to higher emissions for small plants and plants with less leaves. The idea behind this is that leaves keep the released pollen grains within the canopy. These parameters are left out in S13.

Resuspension is considered in H_{orig} only.

H_{orig} and S13 use a total amount of pollen that can be produced per season. In S13, this number defines the end of the pollen season. In H_{orig} it can shorten the pollen season if the model runs out of pollen before the prescribed end of the

pollen season after 30 days. Since this feature is unwanted, the total amount of pollen has been removed in H_{opt}. The two main problems associated with the total amount of pollen are (i) the actual number is basically unknown, and (ii) for some plants (such as birch trees) this number varies considerably by year (years with high total amounts of pollen are referred to as "mast years").

Some plants (such as birch trees) produce less pollen when they grow at higher altitudes (e.g. Sveinbjörnsson et al., 1996; Gehrig and Peeters, 2000). To represent this fact in the model, a reducing factor, $f_{Q,alt}$, is introduced as a function of altitude in H_{opt}. This factor is plant-specific and takes into account the influence of the changing climatic conditions with altitude on the plant.

4 Development of an emission parameterization for pollen grains

The emission process varies from species to species. For example, some grasses need high relative humidities for the opening of their anthers since they have to swell in order to crack. By contrast, low relative humidities create favorable conditions for the release of birch pollen, since birch pollen's anthers open when they are dry (e.g. Fuckerieder, 1976; Puls, 1985; Keijzer et al., 1996; Matsui et al., 1999; Dahl et al., 2013). Since all of the emission parameterizations mentioned in Sect. 3 have been developed for tree pollen, the question remains as to whether they are suitable for herbaceous plants like grasses or ragweed. Since both species are allergenic and responsible for important pollen seasons, it is desirable to have an emission parameterization that is appropriate for these plants as well. Since the model system COSMO-ART is operationally used at MeteoSwiss, we have looked more closely at the two emission parameterizations associated with COSMO-ART (H_{orig} and H_{opt}). The parameterizations show the following disadvantages.

Two plant-dependent parameters in the emission formula are inversely proportional to the emission flux: the leaf area index and the height of the plant. For trees, the variability of these parameters lies in a range where their influence is relatively low. For plants with mean heights of 1 m or less and low leaf area indices, however, these plant-dependent parameters become prominent. Moreover, their use only makes sense if their specific values vary over the model domain or time. Otherwise, they can simply be incorporated into a tuning factor. An important problem in this respect is that the true values of both parameters cannot be known in real time for the entire model domain. Therefore, they are not suited to be incorporated into the parameterization. Additionally, the sensitivity of the simulated emission flux on the height of the plant shows the undesirable feature that small plants emit more pollen than big plants.

Meteorological influences on the emission flux are incorporated into the parameterization via a threshold friction velocity (see Sect. 3.2). Only if the friction velocity reaches this threshold is emission allowed in the model. The threshold value is not constant but includes functions of temperature, relative humidity, and wind speed. Due to the nature of the parameterization, these functions cannot be determined empirically. In our opinion, one of the main drawbacks of this emission parameterization is the mixing of biological and meteorological factors that influence different aspects of the emission process. This leads to complicated formulations that cannot be easily validated via straightforward experiments.

Consequently, we have developed a new emission parameterization, EMPOL, which is based on the biological and physical processes leading to pollen emission. It can easily be adapted to different plant species by adjusting just a few key factors. Plant-dependent values that are variable over the model domain, such as the actual height of plants, are omitted. The meteorological influences are designed in a way that allows simple experimental determination of the corresponding functions.

The description of the pollen season is not part of EMPOL but is taken from the seasonal model developed for H_{opt} (compare Sect. 3.1). It is read into the model as an input parameter.

4.1 Basic concepts

Given that the pollen season has started (in other words: plants are ready to release pollen), the emission of pollen can be seen as a two-step process: first, changes in the meteorological conditions lead to a rupture of the pollen sacs (anthesis). Pollen grains are released from the flowers. Second, the pollen grains that are now exposed to air motions can be entrained into the atmosphere. We have adopted this view by dividing our emission parameterization into two parts:

1. Depending on biological and meteorological conditions, a certain number of pollen grains are released from the flowers and fill up a pollen reservoir. Botanically, this process is called pollen presentation.

2. If meteorological conditions are favorable, pollen in the reservoir are entrained into the atmosphere.

Figuratively, the pollen reservoir can be seen as a surface where the pollen grains rest after being released from the anthers, e.g., a leaf. Such a process has been described by Bianchi et al. (1959) for ragweed; most of the pollen first fall onto the foliage below the flowers before being entrained into the atmosphere. However, this descriptive view should not be exclusive: if the conditions are favorable, pollen grains can be released from the flowers and entrained directly into the atmosphere. In that case, the "reservoir" should only be seen as a way to describe the fact that pollen grains have to be made available before they can be carried into the atmosphere.

Conceptually, the parameterization can be described as follows: a constant factor, $Q_{\text{pollen,day}}$, gives the maximal daily amount of pollen that could be released at the height of the pollen season on one square meter if the conditions for pollen release and entrainment were perfect. All other factors take values in the range between 0 and 1 and describe resistances to the pollen release. These factors include, e.g., unfavorable meteorological conditions, but also a low plant coverage, or an early or late date in the pollen season.

The reservoir R_{pollen} is made up of the pollen that are released from the flowers at that very time step (ΔR_{pollen}) and the pollen that were left in the reservoir after the previous time step ($R_{\text{pollen,old}}$). We assume that the pollen in the reservoir can be lost due to random processes (such as animals brushing against the plant and causing pollen to fall to the ground). This loss is described in the factor Ψ_{random}. Additionally, rain washes out a specific portion of the reservoir which is described via the function Ψ_{precip}. Combining the factors described above, the content of the reservoir is given by

$$R_{\text{pollen}} = \left(\Psi_{\text{random}} \cdot R_{\text{pollen,old}} + \Delta R_{\text{pollen}} \right) \cdot \Psi_{\text{precip}}. \tag{1}$$

The amount of pollen that the flowers release into the reservoir at each time step is given by

$$\Delta R_{\text{pollen}} = \Phi_{\text{plant}} \cdot \Phi_{\text{met}} \cdot \Phi_{\text{biol}}. \tag{2}$$

The factor Φ_{plant} combines the plant-specific variables that define the amount of pollen that could be released per time step under perfect meteorological conditions at a given grid point. It consists of the figure $Q_{\text{pollen},\Delta t}$ (calculated from $Q_{\text{pollen,day}}$) describing the maximum amount of pollen that could be released per time step and per square meter if the grid box was totally covered with the specific plant in the perfect growing state. This maximum number is reduced by factors describing the percentage of ground actually covered with the specific plant ($f_{Q,\text{cov}}[0, 1]$), the course of the pollen season using a mathematical description $f_{Q,\text{seas}}[0, 1]$ (see Sect. 3.1), and the influence of the altitude on the productivity of the plants $f_{Q,\text{alt}}[0, 1]$ (see Sect. 3.3):

$$\Phi_{\text{plant}} = Q_{\text{pollen},\Delta t} \cdot f_{Q,\text{cov}} \cdot f_{Q,\text{seas}} \cdot f_{Q,\text{alt}}. \tag{3}$$

The meteorological influences on pollen release are described via mathematical functions:

$$\Phi_{\text{met}} = f_{R,T} \cdot f_{R,\text{RH}}. \tag{4}$$

These equations describe the probability that anthesis will happen under the current meteorological conditions. Up to now, only temperature and relative humidity have been considered. Since the processes leading to pollen release are slightly different for different plant species, these functions are plant-dependent and need to be adapted for the different species.

Additionally, a switch, Φ_{biol}, is introduced that turns off pollen release as soon as a certain daily amount of released pollen is reached. Under optimal conditions, all ripe pollen could potentially be released before the end of the day. Once the flowers have run out of ripe pollen grains, pollen release will stop even if good conditions continue.

The second part of the emission parameterization describes the entrainment of the pollen from the reservoir into the atmosphere. This process is mainly driven by meteorological conditions, namely moisture and turbulence. The pollen flux is given by

$$F_{E,\text{pollen}} = \frac{R_{\text{pollen}}}{\Delta t} \cdot f_{E,\text{TKE}} \cdot f_{E,\text{RH}}, \tag{5}$$

with Δt being the time step of the simulation. It is assumed that the pollen are dispersed instantaneously and homogeneously within the grid box. The pollen concentration mainly depends on the amount of turbulence that lifts the pollen into the air. This is considered through the function $f_{E,\text{TKE}}[0, 1]$. In moist conditions, entrainment is reduced as pollen grains tend to stick to the surface. The strength of this constraint – $f_{E,\text{RH}}[0, 1]$ – is given as a function of the relative humidity. If pollen grains are released from the anthers in an explosive manner (compare Rohwer, 1993) and entrained into the atmosphere regardless of the turbulent kinetic energy and relative humidity, these two functions $f_{E,\text{TKE}}$ and $f_{E,\text{RH}}$ can be set to a fixed value of one. In that case, the pollen reservoir will not be filled, but all pollen released from the anthers will be entrained into the atmosphere directly.

A flowchart in the Supplement shows the different steps of the parameterization and the influences for each of these steps. If EMPOL is used for other pollen species, the plant-specific values in the following parameters need to be adapted: $Q_{\text{pollen,day}}$, $Q_{\text{pollen},\Delta t}$, $f_{R,T}$, $f_{R,\text{RH}}$, $f_{E,\text{TKE}}$, $f_{E,\text{RH}}$, Ψ_{random}, and Ψ_{precip}. Additionally, the following input fields/values need to be provided: $f_{Q,\text{cov}}$, $f_{Q,\text{seas}}$, $f_{Q,\text{alt}}$, the pollen diameter, and the pollen density.

4.2 Tuning of the emission parameterization

One of the ideas behind EMPOL is the possibility to deduce the main parameters via dedicated experiments (e.g. Michel et al., 2012). In a laboratory, it should be possible to measure the functions relating the amount of released/entrained pollen and specific environmental conditions (e.g. temperature) while keeping the remaining parameters constant. For lack of such experimental data, we had to formulate the missing functions in the parameterization on the basis of measured pollen concentrations. EMPOL contains several parameters that have to be tuned: $Q_{\text{pollen,day}}$, $Q_{\text{pollen},\Delta t}$, $f_{R,T}$, $f_{R,\text{RH}}$, $f_{E,\text{TKE}}$, $f_{E,\text{RH}}$, Ψ_{random} and Ψ_{precip}. We used bihourly birch pollen measurements in Switzerland and simulations with COSMO-ART to derive a first guess for each of these parameters. Taking this version of the parameterization, we simulated two months during the birch pollen season
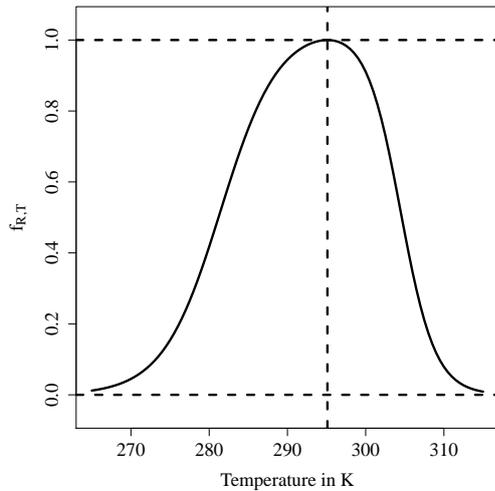
Fig. 1. Function describing the influence of temperature on the release of birch pollen.
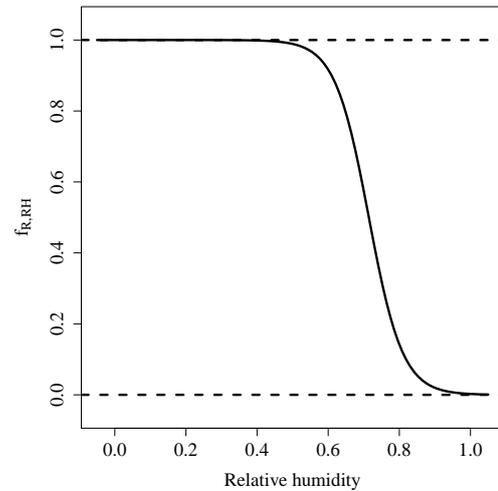

Fig. 2. Function describing the influence of relative humidity on the release of birch pollen.

(April 2010 and April 2011) to improve this first guess. It should be noted that – using birch pollen measurements as a basis – the formulations and values described here are valid only for birch. It should also be kept in mind that using the full modeling system and measured pollen concentrations at this stage introduces some strong assumptions. The most important of these is that any atmospheric transport and dispersion processes are disregarded, i.e., larger pollen concentrations are solely due to larger emissions. The present exercise, therefore, is not more than a "second guess" and still leaves room for improvement based on a true parameter tuning exercise in which the error is minimized by varying all the parameters; this will be performed as soon as corresponding data are available. The present tuning only attempts to render the resulting emissions in a broadly reasonable range, while attempting to describe the physical and biological overall performance.

The parameter $Q_{\mathrm{pollen,day}}$ reflects the overall level of the pollen flux. It was tuned by calculating the overall bias between measurements and model. Based on these values, $Q_{\mathrm{pollen,day}}$ was set to 2.133×10^9 pollen per square meter and per day. The amount of pollen that can be released per time step is calculated from $Q_{\mathrm{pollen,day}}$ considering the following requirements/assumptions:

– Under optimal conditions for pollen release, the flowers can run out of pollen grains before the end of the day.

– The daily cycle of pollen release is not prescribed in the model. The amount of released pollen does not depend on the time of the day but results from the functions $f_{R,T}$ and $f_{R,\mathrm{RH}}$.

In our implementation, we assume that the flowers will run out of pollen grains after 16 h of constant pollen release. $Q_{\mathrm{pollen},\Delta t}$ can then be calculated by dividing $Q_{\mathrm{pollen,day}}$ by the number of time steps during a 16 h-period:

$$Q_{\mathrm{pollen},\Delta t} = \frac{Q_{\mathrm{pollen,day}} \cdot \Delta t}{16 \cdot 3600}. \tag{6}$$

The functions $f_{R,T}$, $f_{R,\mathrm{RH}}$ and $f_{E,\mathrm{TKE}}$ were tuned by calculating the absolute error between the modeled and the measured pollen concentrations for each measuring station. These errors were plotted against the different meteorological variables. These plots were then used to adjust the functions describing the meteorological influences on pollen release/entrainment. The resulting functions are (the corresponding curves can be found in Figs. 1, 2 and 3)

$$f_{R,T} = \frac{1.04}{\left(1 + e^{-0.27T + 76}\right) \cdot \left(1 + e^{0.45T - 137}\right)}, \tag{7}$$

$$f_{R,\mathrm{RH}} = \frac{1}{1 + e^{21\mathrm{rh} - 15}}, \tag{8}$$

$$f_{E,\mathrm{TKE}} = \frac{1}{1 + e^{-2.1\mathrm{TKE} + 4}} - 0.017. \tag{9}$$

In these functions, T denotes temperature, rh the relative humidity, and TKE the turbulent kinetic energy on the lowest model level.

The number of pollen that are lost from the reservoir due to random processes is derived by using the concept of a half-life. Under the premise that only the random losses are effective, we assume a continuous removal from the reservoir such that after 12 h, half of the pollen present in the reservoir is lost. This assumption is used to calculate the percentage of the pollen in the reservoir that are lost per time step due to random processes:
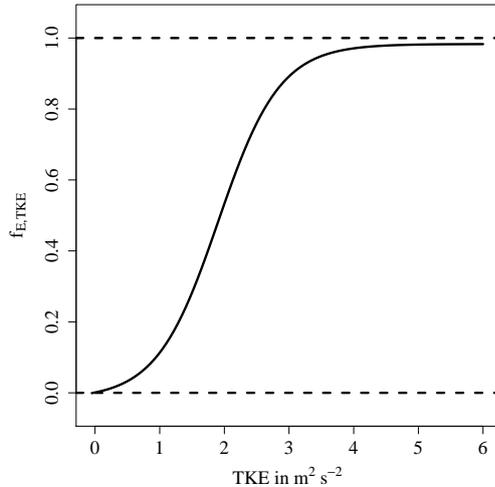
Fig. 3. Function describing the influence of turbulence on the entrainment of birch pollen grains into the atmosphere.

$$\Psi_{\text{random}} = \exp\left\{\frac{\ln 0.5 \cdot \Delta t}{12 \cdot 3600}\right\}, \tag{10}$$

where Δt is the time step of the model given in seconds.

The washout of the reservoir is given by:

$$\Psi_{\text{precip}} = -2000 \, p + 1 \qquad 0 < p < 0.0005, \tag{11a}$$

$$\Psi_{\text{precip}} = 0 \qquad\qquad p \geq 0.0005. \tag{11b}$$

Wherein p denotes the sum of convective and grid-scale precipitation in the model. It is given in $\text{kg}\,\text{m}^{-2}\,\text{s}^{-1}$.

The tendency of the pollen to stick to the surface under moist conditions is given as a function of the relative humidity rh:

$$f_{E,\text{RH}} = 1 \qquad\qquad \text{rh} \leq 90\,\%, \tag{12a}$$

$$f_{E,\text{RH}} = 0.5 \qquad 90\,\% < \text{rh} \leq 95\,\%, \tag{12b}$$

$$f_{E,\text{RH}} = 0 \qquad\quad 95\,\% < \text{rh} \leq 100\,\%. \tag{12c}$$

5 Testing the new parameterization

After debiasing the model using Swiss pollen data from 2010 and 2011, we have simulated the birch pollen season of 2012 using two model configurations. Both use an identical setup, with the emission parameterization being the only difference. First, simulations were carried out using the parameterization H_{opt} (see Sect. 3). Second, we have adapted COSMO-ART to run with our newly developed emission parameterization EMPOL.

5.1 Setup of the simulations

We have used the COSMO-ART version 2.0 in combination with the COSMO version 4.19. The operational COSMO-ART domain of MeteoSwiss covers a large part of central and western Europe, reaching from Portugal in the west to the Balkans in the east, and from southern Italy in the south to the southern parts of Scandinavia in the north. The model is run at a spatial resolution of $0.06°(\approx 6.6\,\text{km})$ with 60 vertical levels and a time step of $60\,\text{s}$. Our simulations start on 21 March 2012 and end on 16 May 2012, covering the entire birch pollen season in central Europe. Every 72 h, a new run is initialized using updated meteorological boundary and initial conditions from the operational NWP modeling system.

We use the birch distribution data set that was produced by Pauling et al. (2012) as the map of possible pollen sources. Pauling et al. (2012) present a method to create plant distribution data sets using birch as an example. This method combines forest inventory and land use data from Switzerland to derive a distribution map. The transfer to a larger domain (southern and central Europe) is achieved by using the Global Land Cover 2000 data set in combination with pollen data.

The mathematical description of the pollen season $f_{Q,\text{seas}}$ is taken from the phenological model developed for the operational numerical pollen forecasts at MeteoSwiss (compare Sect. 3.1). It is used as an input parameter for the emission parameterization in both model configurations.

5.2 Comparison to pollen measurements

The simulated pollen concentrations (mean over 9 grid points) are compared to daily measurements at 34 observational sites throughout Europe for the entire pollen season of 2012. A list of the sites, including their geographical location, can be found in Table 1. Pollen grains were sampled using Hirst type traps (Hirst, 1952).

The level of the simulated pollen concentrations at the start and end of the pollen season strongly depends on the seasonal description ($f_{Q,\text{seas}}$, Sect. 3.1). Since the aim of our exercise is the comparison of the emission parameterizations and explicitly not that of the seasonal description, it has to be made sure that the day-to-day differences between modeled and observed pollen concentrations can be attributed mainly to the emission parameterization. Therefore, we exclude days outside the main pollen season from the exercise. The main pollen season is defined as the period between the first and last occurrence of 70 pollen per cubic meter in the observations (daily means). This corresponds to the pollen class "strong" (compare Table 2). Secondly, we exclude days at the beginning and end of the simulated pollen season, since the upswing and downswing of the simulated pollen season is not yet very well defined. To ensure that these days are not taken into account, we exclude days where the value of the modeled pollen season $f_{Q,\text{seas}}[0, 1]$ is below 0.3.

Table 1. Sites of the pollen measurements and their geographical locations, numbers refer to the numbering in the figures and tables of Sect. 6.

	Country	Town	Lon	Lat
1	Austria	Rosalia	16.3033	47.7030
2	Austria	Vienna	16.3561	48.2488
3	Belgium	Brussels	4.3500	50.8333
4	Belgium	Genk	5.5000	50.9500
5	Switzerland	Basel	7.5830	47.5638
6	Switzerland	Bern	7.4211	46.9477
7	Switzerland	Geneva	6.1500	46.2166
8	Switzerland	La-Chaux-de-Fonds	6.8333	47.1144
9	Switzerland	Lausanne	6.6500	46.5333
10	Switzerland	Lucerne	8.2833	47.0500
11	Switzerland	Münsterlingen	9.2333	47.6333
12	Switzerland	Neuchâtel	6.9166	46.9833
13	Switzerland	Zurich	8.5500	47.3833
14	Germany	Berlin	13.4166	52.5333
15	Germany	Delmenhorst	8.6333	53.0500
16	Denmark	Copenhagen	12.5500	55.6833
17	France	Chalon-sur-Saône	4.8369	46.7933
18	France	Chambéry	5.9169	45.5652
19	France	Metz	6.1822	49.0925
20	France	Montluçon	2.6050	46.3400
21	Croatia	Zagreb	15.9833	45.8166
22	Hungary	Györ	17.6000	47.6666
23	Hungary	Kecskemét	19.6666	46.9166
24	Hungary	Budapest	19.0666	47.5000
25	Hungary	Salgótarján	19.9008	48.0455
26	Hungary	Szolnok	20.2000	47.1666
27	Hungary	Tata	18.3180	47.6388
28	Hungary	Veszprém	17.9163	47.1000
29	Italy	Busto Arsizio	8.8541	45.6119
30	Italy	Legnano	8.9088	45.5969
31	Italy	Parma	10.3333	44.7167
32	Poland	Lodz	19.4650	51.7661
33	Poland	Poznan Morasko	16.9166	52.4500
34	Poland	Poznan	16.8833	52.4166

The observational data has been checked for outliers. Additionally, we excluded sites for which there were less than 7 days of observations within the chosen period (see above).

6 Performance of the different model versions

6.1 Statistical measures

Manual operational pollen forecasts are usually done for pollen classes that reflect more or less the strength of the allergenic symptoms that are induced by the given pollen concentration. These thresholds depend on the sensitization rates and are not constant over regions (Jaeger, 2011). Table 2 gives the thresholds for birch pollen concentrations used for the operational pollen forecasts at MeteoSwiss (Gehrig et al., 2013). Taking this classification, and using a lower and upper limit, the continuous values of pollen concentrations were

Table 2. Pollen classes for birch pollen concentrations (C_{pollen}) used for the operational pollen forecasts at MeteoSwiss (Gehrig et al., 2013).

Pollen class	Pollen concentration in m^{-3}
low pollen load	$C_{pollen} < 10$
moderate pollen load	$10 \leq C_{pollen} < 70$
strong pollen load	$70 \leq C_{pollen} < 300$
very strong pollen load	$300 \leq C_{pollen}$

Table 3. 2×2 contingency table: pairs of measured and simulated values are classified as hits (a), false alarms (b), misses (c) and correct negatives (d).

		Observation	
		yes	no
Forecast	yes	a	b
	no	c	d

classified for both the modeled and the measured values. For each pollen class, a 2×2 contingency table (see Table 3) was completed, taking the occurrence of the desired class as an event and the occurrence of any other class as a non-event. Whether the modeled pollen class is higher or lower than the observed one –and, if it is higher or lower, by how much – cannot be taken into account. Additionally, it has to be noted that the occurrence of any non-event in measurements and simulations is counted as a correct negative. Since a non-event is any pollen class other than the one currently under observation, a correct negative does not necessarily mean that the correct pollen class has been forecast.

In addition to this approach based on the usual manner of manual pollen forecasts, we classified the pollen data using single thresholds as it is usually done for precipitation. This reflects the fact that for most allergenic patients, a pollen concentration in excess of a certain personal threshold triggers symptoms, and thus necessitates the intake of medication. Whether this threshold is just reached or strongly overpassed is of minor importance.

Based on the four numbers in a 2×2 contingency table (compare Table 3), a series of skill scores has been computed (Wilks, 2006). Generally, several of these scores should be looked at together in order to form a complete picture of the performance of a model. We have chosen the Threat Score, False Alarm Rate, as well as the Pierce Skill Score.

The Threat Score, TS, (Eq. 13) measures the proportion of correct forecasts without taking into account the correct non-events:

$$\text{TS} = \frac{a}{a+b+c}. \tag{13}$$

Usually, this is done for rare events where correct non-events are meaningless. In our case, disregarding the correct

non-events is especially wanted since they are not necessarily correct, as described above. TS ranges between 0 and 1, with 1 being a perfect score.

The False Alarm Ratio, FAR, (Eq. 14) gives the fraction of simulated events that were not observed:

$$\text{FAR} = \frac{b}{a+b}. \tag{14}$$

FAR takes values between 0 and 1, and a perfect score renders a value of 0.

The Pierce Skill Score, PSS, (Eq. 15) is a measure that describes the performance of a model compared to an unbiased random forecast:

$$\text{PSS} = \frac{ad - bc}{(a+c)(b+d)}. \tag{15}$$

PSS can take values between -1 and 1. Forecasts worse than a random forecast render negative values. Forecasts better than the random forecast result in positive values. Rare events that are correctly forecast count more than frequent events.

Additionally, we calculated some statistical measures that reflect the pollen concentrations as opposed to pollen classes. This overcomes the problem of classification: values close to the thresholds might lead to a wrong class even though they are not far away from the observed value. By contrast, classes that cover a large interval of concentrations lead to events being counted as correct even if the real value is several factors wrong. We have calculated the correlation coefficient (r^2), the p-value, the root mean square error (rmse), the index of agreement (d_1), the fraction of predictions within a factor of two of observations (FAC2), and the geometric mean bias (GMB) (GAW Report No. 181, 2008).

We have calculated the p-value for r^2. It gives the probability of obtaining the observed data if the null hypothesis is true. In our case, the null hypothesis is: "Observations and modeled simulations do not correlate." If the p-value falls below a predetermined significance level (we use a value of 0.05), one can assume that the null hypothesis is wrong. It has to be noted that the p-value is not a proof that the alternative hypothesis is correct. However, if the p-value is above the chosen significance level, it is not justified to reject the null hypothesis.

The index of agreement (Willmott et al., 1985) is based on the sums of the absolute values of the errors between observations and modeled values. It can take values between zero and one, with one being a perfect score. It is given by

$$d_1 = 1 - \frac{\sum_{i=1}^{N} |P_i - O_i|}{\sum_{i=1}^{N} \left(|P_i - \overline{O}| + |O_i - \overline{O}| \right)}, \tag{16}$$

where P_i denotes the simulated values, O_i the observations, $\overline{O}$ the observational mean, and N the number of data points.

Table 4. Threat Score (TS) and False Alarm Ratio (FAR) for pollen classes using a single threshold. The threshold is given in pollen per cubic meter.

Score	Threshold	H_{opt}	EMPOL
TS	10	0.91	0.95
TS	70	0.65	0.69
TS	300	0.34	0.48
FAR	10	0.03	0.02
FAR	70	0.27	0.21
FAR	300	0.58	0.29

FAC2 gives the fraction of the predictions that are within a factor of two of the observations. It is calculated as

$$\text{FAC2} = \frac{1}{N} \sum_{i=1}^{N} n_i, \tag{17}$$

with $n_i = 1$ if the modeled value lies within a factor of two of the observed value. Otherwise, n_i is zero. A "perfect" forecast (with respect to the factor of two) renders a score of one.

The geometric mean bias is given as

$$\text{GMB} = \exp\left(\frac{1}{N} \sum_{i=1}^{N} \ln P_i - \frac{1}{N} \sum_{i=1}^{N} \ln O_i \right). \tag{18}$$

It has to be noted that GMB is sensitive to the relative error. For example, an observed value of 1 and a simulated value of 0.1 render the same GMB as an observed value of 100 and a simulated value of 10. In reality, the former is of little importance to allergic people whereas the latter is crucial. In addition, the measuring systems cannot distinguish between 0.1 and 1, although they are clearly sensitive to the difference between 10 and 100. Furthermore, in the observed pollen concentrations, values of 0 pollen per cubic meter can occur. This obviously results in an error for the logarithm. Hence, the logarithm of pollen concentrations smaller than 20.1 pollen per cubic meter was set to 3.

6.2 Results regarding pollen classes

We have studied the ability of the model versions to forecast a certain pollen class using TS and FAR. The results using a double threshold to define a pollen class (upper and lower bound) can be found in Figs. 4 and 5. Both model versions show that the ability to correctly forecast a pollen class is different for the different pollen classes. Apparently, it is easier to predict higher pollen classes. This is desirable since higher pollen classes induce stronger allergenic symptoms. In any case, the class "low pollen load" scores worst. This can partly be explained by the fact that the higher pollen classes cover a wider range of pollen concentrations (see Table 2): e.g., for the "very strong" pollen class, the concentration needs to be anything above 300 pollen per cubic meter, whereas for the class "low" the concentration needs to be below 10 pollen
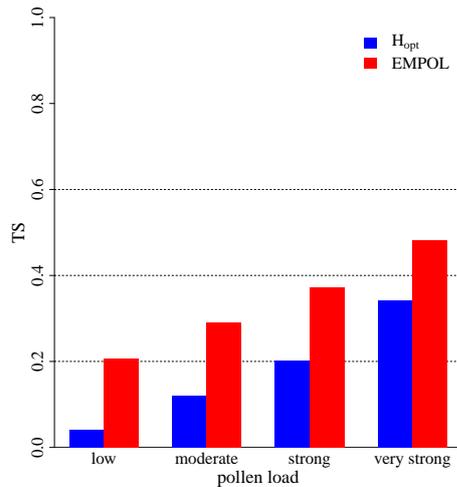
Fig. 4. Threat Score, based on classes defined by an upper and lower limit.
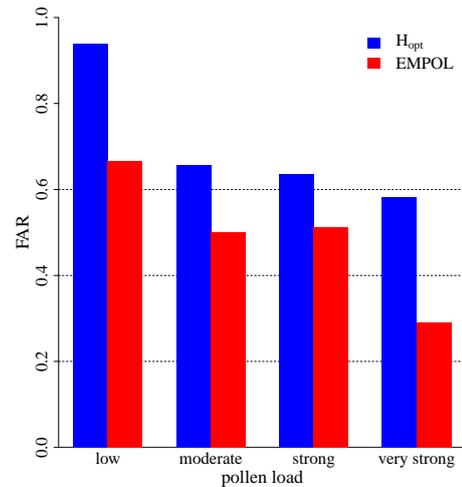

Fig. 5. False Alarm Ratio, based on classes defined by an upper and lower limit.

per cubic meter. Obviously, the latter is more difficult to hit. It has to be noted that the class "low pollen load" is very rare during the main pollen season, both in observations and in the models. The scores are therefore based on only a few data points where the observations fall into this class. Additionally, measurements of small pollen concentrations are uncertain due to the nature of the measuring system.

The same scores have also been computed using a single threshold to define a pollen class rather than an upper and lower limit. The results given in Table 4 still show strong differences between the classes. However, the conclusion drawn from the exercise is reversed: now, the results indicate that it is easier to forecast the lower pollen classes. This demonstrates that the scores are very sensitive to the way an "event" or "non-event" is defined (single or double threshold). Which definition is appropriate depends very much on the question of the study.

The PSS using a double threshold is shown in Fig. 6. For the three lower pollen classes, the values of PSS are on the order of 0.3 (EMPOL) and less than 0.1 (H_{opt}). The scores for the "very strong" pollen class are better for both parameterizations: 0.49 and 0.24, respectively. The results indicate that, indeed, the highest pollen class is better predicted than the lower pollen classes. This is true for both emission parameterizations and for both definitions of an "event" (not shown).

In order to get a sense of the quality of our results, we have compared the PSS values for pollen to the results obtained for operational COSMO forecasts of precipitation at MeteoSwiss. Since precipitation, in contrast to birch pollen concentrations, is not limited to only one season, the analysis has been done for all four seasons. For precipitation, 8 different thresholds are used for the classification, starting with $0.1\,\mathrm{kg\,m^{-2}}$ and ending with $50\,\mathrm{kg\,m^{-2}}$. The corresponding
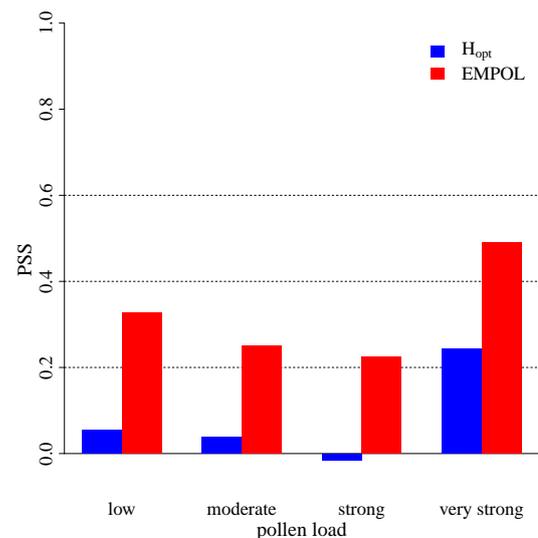

Fig. 6. Pierce Skill Score, based on classes defined by an upper and lower limit.

values can be found in Table 5. Unlike pollen classes, precipitation is better forecast for the lower thresholds. The best score for pollen classes (EMPOL, class "very strong", see Fig. 6) is on the order of the best scores for precipitation (0.6). Likewise, the worst scores for precipitation and pollen (H_{opt}, classes "low", "moderate" and "strong") are in the same range and less than 0.1. Considering that precipitation has already been forecast in NWP models for some decades, this is an encouraging outcome for simulations of pollen concentrations that have been introduced to NWP systems just recently.

Table 5. Values of the Pierce Skill Score for operational COSMO forecasts of precipitation at MeteoSwiss using different thresholds. The four seasons are evaluated separately.

Threshold in $\mathrm{kg\,m^{-2}}$	Spring	Summer	Autumn	Winter
0.1	0.64	0.60	0.63	0.61
1	0.62	0.54	0.64	0.64
2	0.55	0.47	0.59	0.59
5	0.45	0.38	0.55	0.53
10	0.36	0.28	0.48	0.41
20	0.21	0.21	0.35	0.21
30	0.15	0.16	0.26	0.09
50	0.03	0.05	0.10	0.08

Summing up the results, EMPOL scores higher than $\mathrm{H_{opt}}$ for any of the pollen classes and for all of the computed skill scores, regardless of the definition of an "event" (see, e.g., Figs. 4 to 6).

6.3 Results regarding pollen concentrations

The mean correlation coefficient is clearly better for EMPOL than for $\mathrm{H_{opt}}$ (see caption of Table 6). However, to reject the null hypothesis ("Modeled and observed concentrations are uncorrelated".), the p-values have to be less than 0.05. Neither the mean p-value of $\mathrm{H_{opt}}$ nor that of EMPOL fall below this threshold (see caption of Table 6). The picture changes when looking at individual measuring sites (see Table 6). For $\mathrm{H_{opt}}$, the p-values are generally higher, with only the values for three stations below 0.05 (stations 4, 16, and 30). These coincide with the three highest correlation coefficients for $\mathrm{H_{opt}}$. By contrast, EMPOL has p-values of less than 0.05 at 17 stations and only a few stations with very high p-values. Again, the high p-values generally coincide with small correlation coefficients (e.g., stations 2, 14, 15, 21, 23, 33, 34) and vice versa. $\mathrm{H_{opt}}$ only performs better at 7 out of the 34 measuring sites. This can be interpreted such that the biological and physical processes are better represented in EMPOL than in $\mathrm{H_{opt}}$. Both parameterizations have been tuned using Swiss pollen data of the previous years. For EMPOL, this results in better correlation coefficients and very low p-values at most of the Swiss pollen stations (numbers 5 to 13). Nevertheless, apparently the chosen parameterization also works well for many stations outside of Switzerland (e.g., stations 3, 16, 18, 19, 24, 26, 27, 29, 30). Such a refinement of the results for the stations included in the tuning process cannot be found for $\mathrm{H_{opt}}$.

The model's ability to forecast a concentration in reasonable proximity to the observed value can be measured using FAC2 (see Fig. 7). It can be seen that EMPOL performs better at 28 of the 34 stations, which leads to a considerably higher mean value than $\mathrm{H_{opt}}$ has. Overall, about 50 % of EMPOL's simulated concentrations are within a factor of

Table 6. Correlation coefficients (r^2) and their corresponding p-values for the two model configurations at each of the measuring sites. The mean r^2 over all stations is 0.43 for EMPOL and 0.12 for $\mathrm{H_{opt}}$. The corresponding mean p-values are 0.19 for EMPOL and 0.49 for $\mathrm{H_{opt}}$.

	EMPOL		$\mathrm{H_{opt}}$	
Site	r^2	p-value	r^2	p-value
1	0.48	0.03	0.19	0.41
2	0.16	0.54	−0.08	0.76
3	0.80	0.00	0.34	0.21
4	0.35	0.21	0.58	0.02
5	0.74	0.00	0.04	0.87
6	0.76	0.00	0.13	0.59
7	0.85	0.00	0.34	0.22
8	0.41	0.08	0.03	0.90
9	0.72	0.00	0.41	0.13
10	0.86	0.00	−0.02	0.93
11	0.69	0.00	0.20	0.39
12	0.84	0.00	0.17	0.66
13	0.46	0.04	0.25	0.28
14	0.02	0.93	0.03	0.90
15	−0.15	0.69	−0.50	0.17
16	0.88	0.00	0.67	0.02
17	−0.37	0.24	−0.03	0.93
18	0.72	0.01	0.35	0.27
19	0.64	0.06	−0.39	0.29
20	0.43	0.08	0.45	0.06
21	0.16	0.73	−0.27	0.55
22	0.47	0.08	0.38	0.17
23	0.22	0.46	−0.19	0.51
24	0.62	0.01	0.10	0.74
25	−0.29	0.26	−0.05	0.84
26	0.82	0.00	−0.05	0.86
27	0.71	0.08	−0.12	0.79
28	0.31	0.21	0.18	0.47
29	0.66	0.01	0.32	0.27
30	0.67	0.02	0.68	0.01
31	0.39	0.39	0.22	0.64
32	−0.58	0.05	−0.44	0.15
33	0.06	0.84	0.05	0.87
34	0.27	0.34	0.05	0.87

two of the observed values. Using $\mathrm{H_{opt}}$, only about 30 % fall in that range. Similar conclusions can be drawn from d_1 (not shown). Here again, the results from EMPOL tend to be slightly better at the Swiss stations, whereas the results from $\mathrm{H_{opt}}$ do not show this preference (not shown).

The mean rmse is very high for both parameterizations (see Table 7). The individual values for different stations are quite diverse (not shown). In many cases, high values of rmse coincide for the two parameterizations. This suggests that the error could have its origin outside the emission parameterizations. Examples of such external factors are the plant distribution or a wrong altitude of Alpine stations in the model due to the low spatial resolution.
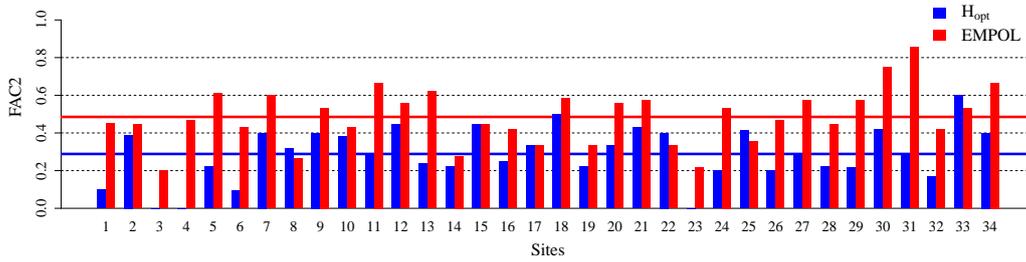
Fig. 7. Fractions of predictions within a factor of two of observations for the two model configurations at each of the measuring sites. The colored lines denote the mean over all stations.
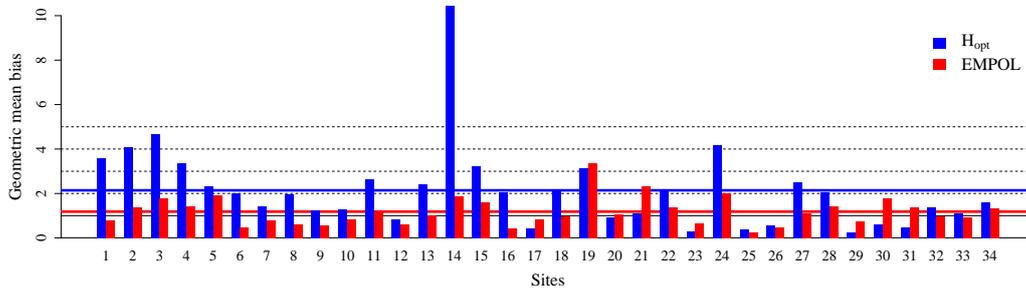

Fig. 8. Geometric mean biases for the two model configurations at each of the measuring sites. The colored lines denote the mean over all stations.

Table 7. Statistical results for the two model configurations based on pollen concentrations. Values are means over all measuring sites. The rmse is given in pollen per cubic meter. Figures in the last column give the number of sites where EMPOL scores better than H_{opt}. The total number is 34 sites.

Statistical measure	H_{opt}	EMPOL	#
d_1	0.31	0.46	29
FAC2	0.29	0.49	28
rmse	585	345	32
GMB	2.14	1.18	26
r^2	0.12	0.43	27
p-value	0.49	0.19	28

A different picture can be found for the GMB (see Fig. 8). Here, again, similar values coincide at some of the stations (e.g., sites 12, 19, 20, 25, 26, 33). However, at the majority of stations, one parameterization clearly shows a stronger bias than the other (e.g., sites 1–4, 14, 21, 24). Here, the difference in the error has to originate in the emission parameterizations since the other influencing parameters are the same for both model versions.

6.4 Sensitivity to mast years

As mentioned before, birch shows a bi-annual variation of the pollen production (years with high pollen production are referred to as "mast years"). The overall level of the pollen concentrations in the model is tuned based on simulations and observations of pollen concentrations of a predefined set of years. Usually, such a set of years will contain both mast years and normal years. This will result in a simulated pollen level somewhere between the level of the mast years and the level of normal years. Therefore, the occurrence or non-occurrence of a mast year in the period chosen for the experiment will have an influence on the performance of the model for the statistical scores that are sensitive to the overall level of the pollen concentrations (e.g., the fraction of predictions within a factor of 2 of observations). Scores that are not sensitive to the overall level (e.g. the correlation coefficient) will not be influenced by the occurrence of a mast year.

We tried to identify the mast and normal years in the period used in this study (2010 to 2012) based on the Swiss observational data. Values of the seasonal pollen index SPI (yearly sum of the observed daily pollen concentrations) for the Swiss observational sites are given in Table 8. It is not possible to make a clear statement. Four of the Swiss observational sites show the highest SPI in the year 2010, three in the year 2011, and two in the year 2012. The variation between the years can be very small (e.g., sites 5 and 9) or very strong (e.g., sites 6 and 11).

Table 8. Seasonal pollen indices for the Swiss observational sites for the years 2010 to 2012.

Site	2010	2011	2012
5	6304	5484	6052
6	9834	16 190	12 156
7	6134	3647	4885
8	3131	1413	3418
9	5816	4426	5839
10	7222	3988	5392
11	5336	13 842	5038
12	4258	2640	3757
13	10 694	12 516	10 246

Looking at these data, it is obvious that it is not possible to make an overall statement about a certain year being a mast year or not even for these stations that are relatively close to each other. The conclusions drawn from the data are not consistent between the different observational sites. Additionally, the difference between mast years and normal years can be strong or very weak, depending on the observational site.

7 Summary and conclusions

We have developed an emission parameterization, EMPOL, for pollen grains for use in NWP systems. EMPOL needs the plant distribution and a description of the pollen season as input parameters. The emission process is separated into two steps. The first step is the release of the pollen from the flowers into a pollen reservoir (pollen presentation). This is driven by temperature and relative humidity. The second step is the entrainment of the pollen from the reservoir into the atmosphere. This is driven by the turbulent kinetic energy. Additional processes that are included: washout of the reservoir by precipitation, loss of pollen in the reservoir due to random processes, sticking of the pollen grains in the reservoir when there is high relative humidity. Under favorable conditions, the available pollen grains can be released rapidly. In that case, pollen release is turned off for the rest of the day. The maximum daily amount of available pollen is a function of the progression of the pollen season.

The meteorological functions have been implemented in a way that allows adaptation for different pollen species. This is important since the opening processes of the anthers are different for different plant species. Additionally, EMPOL uses individual functions for each of the meteorological parameters. This and the separation of the different steps leading to pollen emission facilitate the design of experiments to determine the different meteorological parameters. Another advantage over the parameterizations based on Helbig et al. (2004) (H_{orig} and H_{opt}) is the avoidance of unknown parameters like LAI that introduce unnecessary uncertainty.

Statistical scores show that the model version using EMPOL performs better than the model version using H_{opt}. This is true both for analyses using continuous values and for analyses based on pollen classes.

TS and FAR show that the performance of the model differs for different pollen classes. Using a lower and upper limit to define classes, the higher pollen classes are better predicted than the lower pollen classes. Using a single threshold to discriminate between classes, the lower pollen classes reach better scores. For PSS, the scores of the higher pollen classes are better, both for single and double thresholds. Overall, the scores are comparable to the ones that are reached for operational forecasts of precipitation using the COSMO model. The best score (0.49) is reached for the pollen class "very strong" and EMPOL. The mean correlation coefficient for H_{opt} is 0.12, whereas for EMPOL it reaches a value of 0.43. The results are quite diverse among the different observational sites. Using EMPOL, 49 % of the simulated values deviate less than a factor of two from the observations (H_{opt}: 29 %).

It has to be noted, however, that this parameterization is not yet fully tuned. The functions in the parameterization work in the timescale of a time step. In our case, these are 60 s. It is hardly possible to tune these functions using daily resolved observations. Pollen observations with a higher temporal resolution (bihourly values) have started to be registered on an operational basis only recently. Additionally, it would be desirable to conduct laboratory experiments for the explicit determination of the functions that connect the amount of released/entrained pollen to meteorological variables. Taking these possibilities into account, EMPOL has a great potential to become considerably better in the future.

For the time being, the functions have only been tested for birch pollen emission. However, during the last months, EMPOL has also been employed to parameterize the emission of grasses and ragweed using plant-specific constants (e.g. in the functions for temperature, humidity and TKE). First results are very promising (not shown). Parameterizing the emissions of further plant species, such as hazel, alder, or ash, should follow to enable the operational use of numerical predictions for the main allergenic pollen species.

For the future development of EMPOL, the following paragraph lists some of the possible improvements.

- Conducting field or laboratory experiments to deduce better functions relating meteorological conditions to the different steps of pollen emission.

- Introducing a mechanism that hinders pollen release and/or entrainment for a certain time period after a precipitation event.

- Introducing the influence of rising CO_2 on the pollen production e.g., Ziska and Caulfield, 2000; Rogers et al., 2006). With respect to the present parameterization this could be done in several ways: (1) $Q_{pollen,day}$

could be transformed into a variable field (right now, its value is fixed for the entire domain), (2) a new input field could be introduced that reflects the influence of CO_2, e.g., f_{Q,CO_2}, (3) the influence of CO_2 could even be calculated within the model if COSMO-ART is run in a "full mode" including reactive trace gases.

Supplementary material related to this article is available online at http://www.geosci-model-dev.net/6/1961/2013/gmd-6-1961-2013-supplement.pdf.

References

Bianchi, D. E., Schwemmin, D. J., and Wagner Jr., W. H.: Pollen Release in the common ragweed (*Ambrosia artemisiifolia*), Bot. Gaz., 120, 235–243, 1959.

Dahl, Å., Galán, C., Hajkova, L., Pauling, A., Šikoparija, B., Smith, M., and Vokou, D.: The Onset, Course and Intensity of the Pollen Season, in: Allergenic Pollen: A Review of the Production, Release, Distribution and Health Impacts, edited by: Sofiev, M. and Bergmann, K.-C., Chapter 3, 29–70, Springer Science+Business Media, 2013.

Fuckerieder, K.: Der Graspollengehalt der Luft in Mitteleuropa, Ph. D. thesis, Auswertestelle Aerobiologie des Umweltbundesamtes und Botanisches Institut der Technischen Universität München, 1976.

García-Mozo, H., Galán, C., Belmonte, J., Bermejo, D., Candau, P., Díaz de la Guardia, C., Elvira, B., Gutiérrez, M., Jato, V., Silva, I., Trigo, M. M., Valencia, R., and Chuine, I.: Predicting the start and peak dates of the Poaceae pollen season in Spain using process-based models, Agr. Forest Meteorol., 149, 256–262, 2009.

GAW Report No. 181: Joint Report of COST Action 728 and GURME – Overview of Tools and Methods for Meteorological and Air Pollution Mesoscale Model Evaluation and User Training, 2008.

Gehrig, R. and Peeters, A. G.: Pollen distribution at elevations above 1000 m in Switzerland, Aerobiologia, 16, 69–74, 2000.

Gehrig, R., Pietragalla-Köhler, B., Clot, B., Herren, T., Graber, M.-J., Hauser, M., Moersen, M., and Sallin, C.: Luftpollengehalt in der Schweiz 2012, Tech. Rep. 20, MeteoSchweiz, 2013.

Grewling, Ł., Jackowiak, B., Nowak, M., Uruska, A., and Smith, M.: Variations and trends of birch pollen seasons during 15 years (1996–2010) in relation to weather conditions in Poznań (western Poland), Grana, 51, 280–292, 2012.

Helbig, N., Vogel, B., Vogel, H., and Fiedler, F.: Numerical modelling of pollen dispersion on the regional scale, Aerobiologia, 3, 3–19, 2004.

Hirst, J. M.: An automatic volumetric spore trap, Annals of Applied Biology, 39, 257–265, doi:10.1111/j.1744-7348.1952.tb00904.x, 1952.

Jaeger, S.: The trouble with threshold values for allergy forecasts, in: AEROBIOLOGICAL MONOGRAPHS. Towards a comprehensive vision, edited by: Clot, B., Comtois, P., and Escamilla-Garcia, B., 233–245, MeteoSwiss (CH) and University of Montreal (CA), 2011.

Keijzer, C. J., Leferink-Ten Klooster, H. B., and Reinders, M. C.: The Mechanics of the Grass Flower: Anther Dehiscence and Pollen Shedding in Maize, Ann. Botany, 78, 15–21, doi:10.1006/anbo.1996.0089, 1996.

Marceau, A., Loubet, B., Andrieu, B., Durand, B., Foueillassar, X., and Huber, L.: Modelling diurnal and seasonal patterns of maize pollen emission in relation to meteorological factors, Agr. Forest Meteorol., 151, 11–21, doi:10.1016/j.agrformet.2010.08.012, 2011.

Martin, M. D., Chamecki, M., and Brush, G. S.: Anthesis synchronization and floral morphology determine diurnal patterns of ragweed pollen dispersal, Agr. Forest Meteorol., 150, 1307–1317, doi:10.1016/j.agrformet.2010.06.001, 2010.

Matsui, T., Omasa, K., and Horie, T.: Mechanism of Anther Dehiscence in Rice (Oryza sativa L.), Annals of Botany, 84, 501–506, doi:10.1006/anbo.1999.0943, 1999.

Michel, D., Rotach, M. W., Gehrig, R., and Vogt, R.: On the efficiency and correction of vertically oriented blunt bioaerosol samplers in moving air, Int. J. Biometeorol., 56, 1113–1121, doi:10.1007/s00484-012-0526-x, 2012.

Pasken, R. and Pietrowicz, J. A.: Using dispersion and mesoscale meteorological models to forecast pollen concentrations, Atmos. Environ., 39, 7689–7701, 2005.

Pauling, A., Rotach, M., Gehrig, R., and Clot, B.: A method to derive vegetation distribution maps for pollen dispersion models using birch as an example, Int. J. Biometeorol., 56, 949–958, doi:10.1007/s00484-011-0505-7, 2012.

Puls, K. E.: Scheitert die Pollenflugvorhersage an der Wetterprognose?, Allergologie, 8, 21–25, 1985.

Rogers, C. A., Wayne, P. M., Macklin, E. A., Muilenberg, M. L., Wagner, C. J., Epstein, P. R., and Bazzaz, F. A.: Interaction of the Onset of Spring and Elevated Atmospheric CO_2 on Ragweed (Ambrosia artemisiifolia L.) Pollen Production, Environ. Health Perspect., 114, 865–869, 2006.

Rohwer, J. G.: Moraceae, in: Flowering Plants. Dicotyledons: Magnoliid, Hamamelid and Caryophyllid Families. The Families and Genera of Vascular Plants, edited by: Kubitzki, K., Rohwer, J. G., and Bittrich, V., 2, 438–453, Springer, 1993.

Sarvas, R.: Investigations on the annual cycle of development of forest trees, autumn dormancy and winter dormancy, Commun. Inst. Forest Fenn., 84, 1–101, 1974.

Schueler, S. and Schlünzen, K. H.: Modeling of oak pollen dispersal on the landscape level with a mesoscale atmospheric model, Environ. Model. Assess., 11, 179–194, 2006.

Sofiev, M., Siljamo, P., Ranta, H., Linkosalo, T., Jaeger, S., Rasmussen, A., Rantio-Lehtimaki, A., Severova, E., and Kukkonen, J.: A numerical model of birch pollen emission and dispersion in the atmosphere, description of the emission module, Int. J. Biometeorol., 57, 45–58, doi:10.1007/s00484-012-0532-z, 2013.

Steppeler, J., Doms, G., Schättler, U., Bitzer, H., Gassmann, A., Damrath, U., and Gregoric, G.: Meso gamma scale forecasts using the nonhydrostatic model LM, Meteorol. Atmos. Phys., 82, 75–96, 2002.

Sveinbjörnsson, B., Kauhanen, H., and Nordell, O.: Treeline ecology of mountain birch in the Torneträsk area, Ecol. Bullet., 45, 65–70, 1996.

van Moerbeke, D. (Ed.): European allergy white paper – Allergic diseases as a public health problem in Europe, UCB Institute of Allergy, Braine-l'Alleud, 1997.

Vogel, H., Pauling, A., and Vogel, B.: Numerical simulation of birch pollen dispersion with an operational weather forecast system, Int. J. Biometeorol., 52, 805–814, 2008.

Vogel, B., Vogel, H., Bäumer, D., Bangert, M., Lundgren, K., Rinke, R., and Stanelle, T.: The comprehensive model system COSMO-ART – Radiative impact of aerosol on the state of the atmosphere on the regional scale, Atmos. Chem. Phys., 9, 8661–8680, doi:10.5194/acp-9-8661-2009, 2009.

Wilks, D. S.: Statistical Methods in the Atmospheric Sciences, International geophysics series, vol. 91, 2nd Edn., Academic Press, Amsterdam, 260–276, 2006.

Willmott, C. J., Davis, R. E., Feddema, J. J., Klink, K. M., Legates, D. R., Rowe, C. M., Ackleson, S. G., and O'Donnell, J.: Statistics for the evaluation and comparison of models, J. Geophys. Res., 90, 8995–9005, doi:10.1029/JC090iC05p08995, 1985.

Zink, K., Vogel, H., Vogel, B., Magyar, D., and Kottmeier, C.: Modeling the dispersion of *Ambrosia artemisiifolia* L. pollen with the model system COSMO-ART, Int. J. Biometeorol., 56, 669–680, doi:10.1007/s00484-011-0468-8, 2012.

Ziska, L. H. and Caulfield, F. A.: Rising CO2 and pollen production of common ragweed (Ambrosia artemisiifolia), a known allergy-inducing species: implications for public health, Aust. J. Plant Physiol., 27, 893–898, 2000.