Interactive comment on “The Simulator of the Timing and Magnitude of Pollen Season (STaMPS) model: a pollen production model for regional emission and transport modeling” by T. R. Duhl et al.

Anonymous Referee #1

Received and published: 12 June 2013

The paper is presenting a new model for pollen emission of five tree species and grass made for Southern California. The model is described in details and its application to future-climate conditions are shown and discussed. Pollen modelling and forecasting is gradually attracting more and more attention from various communities, including atmospheric dispersion modellers. Taking into account fundamental problems of modelling the biological processes and difficulties with reliable observational information, the necessity and importance of the study is high. Pollen developments in MEGAN model are also important as the model has the potential for incorporation of the relevant processes and is also widely used worldwide. Therefore, I started reading with high interest. Unfortunately, the impression after finishing was much less bright. I found several methodological problems, which have to be addressed before the paper can be considered for publication. Below, I list my main objections grouping them into sub-sections followed by a list of a few useful references, in addition to those quoted in this and the companion papers.

Regional applicability.

The species considered in the paper are not all “natural” for Mediterranean type of climate. For instance, I cannot imagine that birch model can be developed or verified based on the data from such region. The habitat of this tree is located very far to the north of it. In Southern California, birch is heavily stressed by heat and water availability, so that the parameterizations based on the regional data have nothing in common with actual birch behavior. In particular, the base temperature of 9.1°C suggested in the paper is confusing: typical range suggested in various works is 3-5°C. The same is true for heat sum: typical value reported in the literature is around 100 degree-days (smaller in the north, larger in the south), which has nothing common with the baffling 620 dd suggested in the paper. This problem is also evident from the companion paper, which presents the observation results. Peak concentrations during the season about 5 pollen/m3 is negligibly small (about 1000 times smaller than in the main birch habitats). Therefore, I have to conclude that the birch model parameters are unrealistic and the model is not suitable for the main tree habitat. This is also confirmed by poor model-measurement comparison (discussed below).

As an example of more appropriate species, the realistic model parameters are obtained for e.g. olive, the natural Mediterranean tree. The base temperature of 9.1°C is still very high (usually reported values are below 5°C) but very low threshold of 490 dd compensates the unusually high base temperature value. Thus, for base temperature of 0°C the threshold would be about 1000-1200 dd, i.e. the season timing predictions will be quite comparable.
From the above, it is evident that the model presented in the paper has much narrower applicability than it is claimed. That needs to be corrected and the ambitions scaled down to the actually delivered results. Non-natural species in California should be excluded (first of all, birch). Strict binding to Southern California should be made clear already in the title, abstract, and introduction. I understand that the basic approaches are universal – but also trivial and known for decades (e.g. many references go back 20-30 years). The devil is in details: it is the data existence and availability, as well as the possibilities of generalization of local and regional findings that presently limit the pollen model development worldwide. And from that point of view, current study is strictly South-Californian.

Methodology

The authors accept many values from unconnected studies, often very old ones. This is normal practice in science but still requires care and is outright dangerous in case of pollen: natural variability is extremely high, as well as the sensitivity of the results to the setup of the field and lab studies. For instance, heat-sum threshold is known to vary by a factor of two or even more at a spatial scale of just a few tens of km, especially in complex-terrain regions. Taking a single value for the whole region is much too crude approach.

In several cases the values are extrapolated across species “due to lack of data” without justification and verification. This is not acceptable. The species, for which the data are not available – and again birch is to be mentioned first – should be excluded from consideration.

Pollen counts is a poor type of input data for determining the start of flowering. The authors paid no attention to vast amount of publications analyzing early- and late-season long-range transport (LRT) episodes, which dramatically change timing of the pollen season, i.e. period with substantial pollen concentrations in the air, as compared with local pollen release season, i.e. the flowering period, the goal of the study. The difference can be as large as a month! The impact of LRT episodes is more moderate only for the species native in the area. For taxa with the main habitat outside the region, the pollen season can be almost entirely decided by a few LRT episodes, which have little connection to regional developments. This is the probable reason for poor model performance for several species (as shown in the companion paper). Phenological data should be used instead for more accurate model parametrization.

The authors have excluded the year 2007 without any justification, just because it looked differently from the others. This is quite shocking: such thinning of the datasets should have very strong justification. Actually, strong meteorological variation would rather help to parametrize the model and improve its ability to reproduce the phenological processes under varying external forcing. Existence of such non-trivial year should be considered as the advantage of the study rather than its drawback. How can the model be applied to future climate, where extremes are more probable, if even in the present situation part of the data is excluded at the very beginning?

Credibility of the results and model evaluation

Evaluation of the model is not presented at all. Instead, the reader is referred to another paper, in different journal and not yet accepted for publication. This is the major problem: the presentation of the model is bound to include its assessment. Companion yet-to-be-accepted paper in different journal does not qualify for that. Nevertheless, I have read the companion paper in order to understand how the above-criticized methodological problems affected the performance. Several points are clear: birch is indeed practically not represented in the region. For comparison, typical concentrations in Central and Northern Europe during the main birch pollen season exceed 1000 pollen/m³, maximum going over 20,000-30,000 pollen/m³, whereas in the current application the counts never exceed 10. No surprises that the model failed it. Walnut and mulberry largely follow similar suite: their concentrations are very low and model predictions have little common with observations. As a result, only grass, olive and oak have substantial representation in the region and non-negligible pollen concentrations.
The evaluation is performed for a single 2010 season, which is insufficient for the model with climate-related ambitions. Difficulties with access to pollen observations also exist in Europe but it cannot justify application of untested models for predicting the future climate conditions. I included a few references that showed the climate response in pollen seasons is very complicated. Some species start flowering earlier, others show later season or appear neutral, and in many cases the response is region-dependent. This again stresses the necessity to evaluate the model for a large variety of conditions before making far-reaching conclusions at climate scale. And I again was missing the rainy year 2007 excluded from both parameterization and evaluation. Does it mean that the model fails if? If yes, why should the reader expect it to work for different climate conditions?

Comparison of the model formulations with other models is entirely missing. How does the suggested parameterization meet / contradict / improve the existing models in Europe and the US? Several models are quoted in the companion paper, which includes some discussion. Why was it not done here in a systematic way?

Finally, as seen from the companion paper, the model showed poor performance for the bulk of the considered species — except for olives and, may be, oak. With such scores, I see no way to approach climate studies. It is not possible to discuss 5 days of the shift of the season if the evaluation showed the error of as much as 1.5 months in the season start (e.g., grass).

A few specific comments, largely related to the above:

P.2330 line. 12-15. This is confusing. The TOTAL pollen produced by a tree during specific season is independent from the conditions during that very season. They are entirely controlled by the previous season when the male flowers are formed — as stated later in the paper. I guess, the authors have mixed-up the daily production and total seasonal production, the first one indeed being controlled by actual meteorological conditions. If yes, it should be stated clearly.

P.2331, line.12-13. The so-called sequential model (heat accumulation starts after chill units are all collected) used by the authors is not always the best approach for explaining the flowering time of several trees. In many cases, parallel model with fixed start of heat accumulation has proven to be better. This problem should be at least discussed.

P.2331, line.15. It is a very well known that GDD threshold is a function of location, with its value varying by a factor of times for various parts of the habitat area. Therefore, the value(s) suggested by any specific publication is(are) valid exclusively in the region (possibly, small) around the study place. To the best of my knowledge, no extrapolation algorithm exist, i.e. the thresholds have to be determined by fitting the model output to the data at maximum number of points and interpolation between them has to be done with high care. This is among the biggest challenges of the pollen source terms developments.

P.2331, line.16-20. Pollen counts can be very misleading when determining the start of flowering (see above).

P.2331, line.16-17. Problems with the methodology are implicitly acknowledged by the authors themselves: they excluded 2007 because of rainy end of the season. But it “automatically” recognizes the fact that the model cannot deal with such conditions.

Section 2.1. The authors claim that there is essentially no data for birch to parameterize the model. However, this is the most-studied tree in Europe. I roughly estimate that 30-40% of aerobiological publications are dedicated to it or use it as one of target species.

Section 2.3. It is a well-known fact that many trees have bi-annual cycle of total seasonal pollen release. Why does this model have no trace of it?

Section 3. Before going into the climate simulations, the model must be evaluated properly, which is not done. After reading the companion paper, I had severe problems
believing the conclusions presented in this section. I would drop this section entirely until the model is improved and its ability to reproduce present climate is confirmed by detailed evaluation.

Useful references


Laaidi, M. (2001) Regional variations in the pollen season of Betula in Burgundy: two models for predicting the start of the pollination, Aerobiologia, 17, 247-254


sites in Andalusia and the effect of the expected future climate change, International Journal of Biometeorology, 49, 184-188


Interactive comment on Geosci. Model Dev. Discuss., 6, 2325, 2013.