Author response to Anonymous Referee #1, 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. (Geosci. Model Dev. Discuss., 6, C767–C775, 2013)

We greatly appreciate the detailed and constructive comments by the reviewers and the opportunity to improve the manuscript. We have revised the manuscript extensively based on the suggestions and believe that the new version is significantly improved as a result of the reviewers' contributions. Please note that we have pasted sections of text from the original Referee report (Italicized and in a different font than our responses) containing the suggested revisions or comments and have responded to these with numbered Author comments.

"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 9.1C suggested in the paper is confusing: typical range suggested in various works is 3-5C. 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)."

R1 Author response 1: Pertaining to the birch parameterizations, we have added an explanation of why we chose the base temperature of 9.1°C and the particular sequential model that was selected by adding the following text (Section 2.1.2, lines 300-305 and 338-374):

"The sequential chill-heating model developed for olives by De Melo-Abreu et al. (2004) is used to simulate pollen season for tree species in STaMPS with chilling requirements, since sequential models have been identified as being appropriate for phenological simulations of both birch and olive in similar climates (Jato et al., 2007; De Melo Abreu, 2004)." …

"For olive, walnut and birch species, we use the same optimum and breakpoint chilling temperature values as those selected for olives in De Melo-Abreu et al. (2004), since these values were calculated for a similar climate as the present study, and observations (e.g., Jato et al., 2007; Warmund et al., 2009) suggest that birch and walnut trees have similar optimal chilling temperatures as olives. A threshold chilling quantity of 58 chilling units was selected for walnuts (Warmund et al., 2009); 432 chilling units was assigned to olives (De Melo-Abreu et al., 2004) and was also assigned to birch (for which published values for birch growing in similar climates were not available).

T_B values close to 0°C have been suggested for birch growing in high latitudes although latitudinal gradients in base temperatures as well as chilling requirements have been observed in many tree species including birch (Myking and Heide, 1994) and grasses (Heide, 1994). We tested a range of base temperatures for birch using the Pasadena data but found little difference in terms of percent standard deviation in accumulated GDD on peak birch pollen count dates using a base temperature of 0°C compared with 9.1 for the years included in the analysis. Most studies of *Betula* have been performed on high-latitude ecotypes although Jato et al. (2007) studied and modeled the onset and length of flowering in *B. pendula* and *B. alba* populations in Spain using several modeling approaches and base temperatures and found that model parameterizations developed for olives (Galán et al., 2001b) resulted in the lowest deviations from actual versus predicted peak pollen date for *Betula*. Data regarding optimal T_B values for birch and walnuts growing in climate zones similar to the study domain are sparse therefore, following a similar approach as Jato et al. (2007), we applied a model developed for olives (De Melo-Abreu et al., 2004) to birch and walnut species within the domain. It should be noted that Jato et al. (2007) used the olive model of Galán et al. (2001b), while we have employed the De Melo Abreu et al. (2004) approach which allows for the calculation of devernalization during chilling calculations. De Melo Abreu et al. (2004) evaluated the inclusion of devernalization against the same chilling calculation approach used by Jato et al (2007) and found results to be more accurate, and the model more physiologically meaningful, when devernalization was considered. It has been long known that species with chilling requirements can lose a portion of their accumulated chilling when temperatures exceed some threshold value (Richardson et al., 1974; Gilreath and Buchanan, 1981). The GDD thresholds for these species were determined using the Pasadena pollen counts as described in Section 2.1.1, with GDD accumulation dates each year beginning on the date when chilling requirements were met for each chilling species according to the Pasadena meteorological data."

We would also argue that model performance for birch wasn't poor as suggested by Referee #1. As stated in the companion paper (Zhang et al., BGD, 10, 3977-4023, 2013) the model exhibited reasonable agreement with both birch and mulberry (which are both non-native to the study region) pollen observations as well as for olive in terms of the predicted day-by-day variation trends and mean values. The fact that birch concentrations were low in the study domain reflects the relative scarcity of this species within the region and also reflects the fact that fractional vegetation cover is much lower in the arid study domain compared with the heavily-treed regions found in many higher latitudes. As a matter of comparison, while Referee #1 may be accustomed to seeing pollen counts on the order of several 1000 m^{-3}, even among oak, one of the most prevalent aeroallergens in our study domain, the peak concentrations are generally less than 300 grains m^{-3} (based on the Pasadena dataset used in the study). The *B. pendula* trees occurring in the domain are found only in the urban tree inventories according to the input datasets (i.e. none are present in forests according to the FIA database) for the region, and our research has shown that this species is a popular landscaping tree in S. CA. Additionally, *B. pendula* is not the only birch tree found in the study domain; there is at least one birch species (*B. occidentalis*) whose native range does extend into the study domain (including both California and Nevada; NRCS USDA Plants Database, http://plants.usda.gov/java/). Table 1 now reflects the fact that other birch species besides only *B. pendula* are present in the model domain. We have also expanded the summary of the transport model evaluation results (lines 593-605). Lastly, as discussed in Author Response #5 (below), the results of a sensitivity analysis using and outer domain simulation to provide dynamic boundary conditions for the base case pollen transport run suggested that nearby regional or local pollen sources were most likely dominating the concentration variation of receptor sites.

"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."

R1 Author response 2: We agree with Referee #1 that the parameterizations developed in StaMPS are explicitly for species occurring in the particular climate zones characteristic of the study domain, and we did not intend to convey in the manuscript that the exact model approaches, GDD thresholds, base temperatures, etc. selected for each species should be used to simulate pollen season in climatic regions differing substantially from those represented in the study. We have made numerous changes, including to the abstract, moving the text "in a study domain centered over Southern California (S. CA)" to the first sentence, "climate zones represented" was inserted in line 39, "in the study region" was inserted in line 47, and "in the study region" (line 53). We have also inserted the following text into the Introduction (lines 96-107, also see lines 759-765):

"The relationships between meteorological parameters and pollen season described here are only applicable to this study region (or to study domains in similar climate zones) since observational data (including pollen count data and phenological observations) taken from within the model domain as well as literature-derived relationships from studies conducted in similar climate zones were used to develop the parameterizations. Plans for development of modules for additional species as well as considerations for simulating species occurring in other climate zones are discussed in Section 4.2."

We have endeavored to more fully describe that the STaMPS modeling framework is in fact flexible with respect to simulations in different climate zones because it includes several types of budburst models (thermal time, sequential, and alternating; Section 2.1, lines 160-185), and that in principal, GDD and chilling thresholds and base temperatures obtained from published studies of the selected species growing in other climatic regimes could be applied, and the choice of budburst model modified, in order to perform simulations in other locations (Section 4.2, 825-837). However, it is also noted that for some species, extensive additional phenological and aerobiological datasets would be required to inform and validate simulations performed in other regions (Section 4.2). STaMPS is designed such that modules can be modified and substituted not only for individual Plant Functional Types (PFTs), but also for species occurring in different climatic regions. Plans to incorporate additional climate-zone simulation capabilities as well as new species are discussed briefly in Section 4.2 (lines 822-840).

"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."

R1 Author response 3: We include a lengthy discussion of the observed variability in the various input datasets used in the model (Section 4.1). We acknowledge that the first version of the manuscript excluded some much-needed descriptions of and justifications for the specific choices made regarding which studies/models, forcing threshold values and base temperatures were assigned. We hope that in the revised manuscript it is clearer that we only applied literature-derived relationships between meteorological variables and flowering, as well as and critical forcing and heating thresholds from previously published studies that are relevant to the species in question and representative of the climate zones within the domain (e.g. lines 104-105, 188-190, 300-305, 338-374, 750-752). We also have placed more emphasis on the fact that the selected GDD thresholds are observationally constrained (lines 737, 760-763) which should serve to increase confidence in the model. We acknowledge that these descriptions were deficient in the first version of the manuscript, but in this revised manuscript we have now added significant additional descriptions of and justifications for the approaches used for the species (e.g. the base temperature and modeling approach applied to birch, as described in Author response #1, above, and in Section 2.1.2 of the manuscript) and we have also placed greater emphasis on the fact that the relationships are valid only for simulations performed in similar climate zones (throughout the manuscript as described above in Author Response #2). Please note that Author response 5 is also relevant to these comments and includes additional instances of revisions made in response to them.

"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."

R1 Author response 4: As described in Section 2.1.1 (lines 246-256) and Appendix B, STaMPS does include algorithms for calculating variable heat-sum thresholds and assigns them to species (Quercus) in which they have been observed in similar climates as our study. Table 1 includes the equations used to determine forcing thresholds for oaks.

"In several cases the values are extrapolated across species "due to lack of data" with-out 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."

R1 Author response 5: We hope that in the revised manuscript it is much more clear that instead of simply accepting "many values from unconnected studies" or assigning values "due to a lack of data" we only applied literature-derived relationships between meteorological variables and flowering, as well as and critical forcing and heating thresholds from previously published studies that are relevant to the species in question and representative of the climate zones within the domain. See Author Responses #1-3 (above) for the exact revisions made. We agree that phenological observations are the ideal datasets to inform predictive models of budburst, and many of the studies that we used to parameterize these relationships for the selected species were based on phenological observations (e.g. *Bromus*: Gleichsner, J. A. and Appleby, A. P. Weed Sci., 44, 57-62, 1996; *Quercus*: phenological observations provided courtesy of Dr. Walt Koenig of Cornell University; and Olea: De Melo-Abreu, J. P., Barranco, D., Cordeiro, A. M., Tous, J., Rogado, B. M. and Villalobos, F. J. Agr. Forest Meteorol., 125, 117-127, 2004). We have now included a discussion of the possible contribution to observed pollen from long-range transport (LRT) episodes and the steps we took to consider LRT in our simulations as well as acknowledging the limitations inherent in using observed pollen counts as proxies for dates of peak pollen release (Section 4.1). The following text was inserted into Section 4.1 (Lines 739-765; also see lines 591-592 which reference the use of a nesting domain to characterize boundary conditions):

"Long-range transport of pollen may also complicate the interpretation of observed atmospheric pollen curves (e.g., Siljamo et al., 2008), especially for non-native species within a study domain that could have a much larger presence in regions outside of the domain, such as birch species. Jato et al. (2007) compared *Betula* pollen curves with phenological observations from two species of birch growing in Spain, and found that for the calculation of GDD requirements, using dates of peak pollen concentration yielded similar results as use of phenological observations, although lags between dates of peak flowering and peak pollen count have been reported for a number of species (e.g., Latorre, 1997). Although we cannot exclude the possibility that pollen transported over large ranges may have influenced the pollen count data used for the determination of heat thresholds, a multi-year aerobiological dataset such as that employed in the present study should be dominated by local signals over time, and numerous studies (of the same genera selected for initial simulation in STaMPS) performed in similar climatic regions were consulted during model development. In an effort to include the potential contributions from pollen transported into the domain during the evaluation simulation period, a nested domain scheme in which the 4-km S. CA domain pollen transport simulations used boundary conditions provided by simulation results of a larger 12-km domain that covers the entire states of California and Nevada (Figure 1; as described in Section 2.4) was used; the simulations results indicate contributions from California and Nevada outside the S. CA domain to pollen counts at the observational sites were insignificant (Zhang et al., 2013a). Our study is the first to predict pollen production for multiple species in the western half of the US, and is constrained with observations taken from within the study domain as well as built using numerous previously published relationships for the selected genera occurring in similar climate zones as the domain, therefore STaMPS improves capabilities for predicting pollen season in this region. Nonetheless, these limitations should be kept in mind when interpreting the pollen dispersal results.**"**

Finally, as pointed out in our companion paper (Zhang et al., 2013), although LRT might be expected to affect "spatial distribution and magnitude of pollen concentrations at the local scale (Zink et al., 2011), there should be a correlation between the timing of pollen emission and that of concentration on the regional scale", and using birch performance as an example, the regionally-observed and simulated regionally-averaged mean peak pollen concentration dates both occurred in Mid-May with good performance also seen in mulberry, oak, and walnut. We also did a sensitivity run by adding outer domain simulations to provide dynamic boundary conditions for the base case pollen transport run. The outer domain covers the whole of California and Nevada. It can be treated as a demonstration to test the impact of LRT to the receptor sites in southern California. However, the simulation results only showed slight improvement under favorable wind conditions, which may highlight the importance of nearby regional or local pollen sources to the concentration variation of receptor sites.

"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?"

R1 Author response 6: Perhaps we were not clear in how we used the pollen count data and why 2007 was excluded: we have re-worded the explanation (lines 219-220) for the exclusion of this year from the set of years where cumulative GDD values were calculated and averaged by genus to determine GDD thresholds. Rains were so frequent in the early spring period that for many genera peaks were not observed at all during the expected times (or others). In one of the helpful references in the list provided by referee #1 (Galán et al., 2001), we noted that a similar approach was taken in that study. In addition to hopefully making the exclusion and the reason for it more clear in the revised manuscript, the following text has also been added (lines 220-223):

"In evaluating several models for their ability to predict the start of *Olea* pollen season in Spain, Galán et al. (2001a) excluded two years out of an 18-year olive pollen dataset due to rain events occurring near the beginning of olive pollen season."

"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. "

R1 Author response 7: As is the case for Geoscientific Model Development Discussions, articles accepted to Biogeosciences Discussions (i.e. the journal to which our companion paper was submitted) are citable upon publication, even in the Discussions version of the journals, and journals such as Geoscientific Model Development and Biogeosciences do screen submissions, only publishing those of acceptable quality. Given the breadth of the modeling effort, we knew that more than one manuscript would need to be prepared and submitted in order to publish a satisfactory description of the modeling framework, yet we could not identify a single journal that was appropriate to submit both papers. Therefore we opted to submit to two EGU Copernicus journals, both of which have high standards for quality. There are only two ways to evaluate the output from the STaMPS model: Phenological observations or pollen counts. Since phenological observations were unavailable (except the oak data that was already used to develop the model), comparing model output with aerobiological data is the only remaining option. Simulating the dispersion of modeled pollen produced in STaMPS and available for release into the atmosphere is non-trivial and requires accurate representations of wind fields, mixed boundary layer height, particle physics, etc. One of the key modules to incorporate the daily amount of pollen available for release predicted in STaMPS into the existing air quality simulation platform (WRF-CMAQ) is the pollen emission flux parameterization. In our companion paper (Zhang et al., 2013) we scale the pollen emission rate based on wind effects, which obviously adds more uncertainty to the results. However since the final product of the dispersion model is what is needed for comparison with observations (pollen counts), we decided that the manuscript describing the dispersion model and results was the appropriate platform for discussing model performance and comparison with observed values. Again we suggest that model performance is not poor as Referee #1 asserts. In any case we have expanded the descriptions of model evaluation in Section 2.4 (lines 593-605).

"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."

R1 Author response 8: As mentioned in Author Response 1, and described in Section 2.4 (lines 593-605), we assert that model performance for birch was not poor as suggested by Referee #1. As stated in the companion paper (Zhang et al., BGD, 10, 3977-4023, 2013) the model exhibited reasonable agreement with both birch and mulberry (which are both non-native to the study region) pollen observations as well as for olive. The fact that birch concentrations were low in the study domain reflects the relative scarcity of this species within the region and also reflects the fact that fractional vegetation cover is much lower in the arid study domain compared with the heavily-treed regions found in many higher latitudes. Both mulberry and *B. pendula* trees occurring in the domain are found only in the urban tree inventories according to the input datasets (i.e. none are present in forests according to the FIA database) for the region, and the pollen monitoring station used to infer the peak pollen dates was also in an urban location. Although LRT may have played a role in the observations at certain times, the multi-year nature of the dataset would be expected to be dominated by local signals when viewed as a whole. See Author Responses 1-7 above for the specific changes that have been made to the manuscript to address these concerns.

"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 it? If yes, why should the reader expect it to work for different climate conditions?"

R1 Author response 9: We hope that the revised manuscript now makes it clear that we are not claiming to project climate effects on pollen season in climates outside of our study domain. Author Response 6 addresses the exclusion of 2007. We have added some additional discussion using several of the references that Referee #1 suggested which expands the discussion of our results into comparisons with other studies/regions (see Author response 10, below for specific changes made in the manuscript). Galán et al. (2001, from the list of references suggested by referee #1) used a similar approach in the development of their pollen forecasting model: a number of years worth of observational data were used to constrain the model and a single year was used in the model evaluation. Obviously, more years for comparison is better than fewer, however we had a finite dataset for use during the course of the study and we allocated the data in what we deemed to be the most responsible manner.

"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)."

R1 Author response 10: We have included additional discussion of our approach with respect to other models and have included some comparison of how our results contrast with other model predictions (lines 646, 660-662, 664-670) and observations (lines 643-645, 662-664, 687-689). As mentioned in several previous Author Responses, we disagree that the model evaluation indicated poor performance; with respect to the grass evaluation, one would not expect the pollen counts to agree completely with simulations when just two grass species were included in the model, whereas the observed pollen counts will reflect the contribution from all grass species flowering in an area at any given time.

"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."

R1 Author response 11: We have deleted "in a given season" to avoid confusion since the point of this sentence was to list the variables known to govern pollen production but not the timescales on which they operate. In Section 2.3 of the manuscript we acknowledge that precipitation a full year prior to the start of pollen season in trees can affect the amount of pollen produced and that is in fact how STaMPS calculates the pollen production size for trees.

"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."

R1 Author response 12: We have added discussion of this as described in Author Response 2.

"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."

R1 Author response 13: (same as Author response 4) As described in Section 2.1.1 (lines 246-256) and Appendix B, STaMPS does include algorithms for calculating variable heat-sum thresholds and assigns them to species (Quercus) in which they have been observed in similar climates as our study. Table 1 includes the equations used to determine forcing thresholds for oaks.

"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."

R1 Author response 14: this is dealt with in Author Responses 5 and 6, above

"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."

R1 Author response 15: we agree and we have clarified (lines 297-304, 338-374) in the manuscript that we meant that birch is not well-studied in climates similar to our study domain

"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?"

R1 Author response 16: This effect has not been included in STaMPS; we have added the following text into Section 4.1 (lines 710-712):

"Additionally, some species exhibit bi- or even triennial variations in pollen production (e.g. Celenk et al., 2009), and these effects are not represented in STaMPS."

"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."

R1 Author response 17: This has been addressed in Author Responses 5-10, above

"Useful references"

R1 Author response 18: Thanks for the references; we have included several of them in the revised manuscript (Siljamo et al., 2008; Ziello et al., 2012; Clot, 2003, Galán 2001, Emberlin, 1999; Linkosalo et al., 2008)

Author response to Anonymous Referee #2, 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. (Geosci. Model Dev. Discuss., 6, C1019–C1026, 2013)

"… the publications shows some methodological drawbacks that should be addressed before publication… First, I am missing a section describing the materials used. The data is introduced in the sections where it is used for the first time. It would be easier to have an overview at the beginning of the paper (after the introduction)."

R2 Author Response 1: we have now introduced all of the datasets used as suggested after the Introduction at the beginning of Section 2 (lines 125-137). The following text has been inserted:

"An aerobiological pollen dataset for 2003-2010 collected at the California Institute for Technology (CalTech) campus in Pasadena, CA was used for the selection of species to be included in the initial simulations along with the expertise of co-authors who have been studying local pollen in S. CA for a number of years. Only species with known allergenicities that typically flower in the March-June period in the study area were selected including species that occur naturally throughout the domain (such as oak) and those that occur mainly in urban environments in the domain but are also important allergens in other regions (e.g., birch). The Pasadena pollen data as well as (for oak species) some phenological data were also used for the determination of various threshold values for flowering as described later. Species composition and fractional vegetation cover within the model domain were determined using the datasets presented in Table 2 and described in Appendix A."

"Second, apparently the STaMPS model was validated only indirectly by incorporating its output into a transport model. The resulting pollen concentrations are compared to count data. However, the simulated pollen concentrations not only depend on the output of the STaMPS model, but also on the emission parameterization and the transport/diffusion processes within the transport model. No numbers are given with respect to the accuracy of the STaMPS model regarding current climatic conditions. Cf. eg. Pauling et al. 2013, Aerobiologia, for phenological model performance assessment. In addition, this publication contains scores that can be used for comparison. It would be good to know how well the model can predict the beginning and magnitude of the pollen season. This should be done before using the model with a future climate. Lacking such a validation with current data, I think that it is not justified to make predictions in the future."

R2 Author Response 2: While we agree that additional model evaluation is always desirable, our companion paper evaluates the results of the STaMPS model in the only way possible considering that the sole phenological dataset available to us was used for development of model parameters (and thus couldn't be used to evaluate the model), leaving aerobiological pollen comparisons the only viable option. Although Pauling et al., 2013 did suggest a nice approach for evaluating phenological models, our co-authors also fairly rigorously evaluated the model and as described in our companion paper (Zhang et al., 2013, P. 3992-3993) the performance was encouraging; for instance taking birch, oaks, mulberry, and walnuts as examples the model accurately predicted the regionally-observed peak dates of pollen concentrations and length of pollen season (with the exception of a some early birch counts observed at a couple of the observation sites, which could be attributable to long-range pollen transport, a phenomenon that is now discussed in Section 4.1, lines 739-765). Our companion paper also compared the day-by-day variation trends of the amount pollen available for release predicted by STaMPS to the pollen count data (Figure 4 and Section 3.2 in Zhang et al., BGD, 2013). This is an attempt to evaluate the performance of the results directly from STaMPS to observed value. For mulberry and olive, both the predicted onset and duration of flowering season from STaMPS closely match the corresponding pollen count temporal pattern at all available sites. Also observed in Zhang et al. (2013), e.g. in Fig. 9, is the fact that for most of the nine observational stations used to compared modeled vs. observed mean and maximum concentrations, good agreement was found for all species except walnut concentrations were generally under-estimated and simulated maximum concentrations were too high for grass although mean grass concentrations closely match observations for 6 of 9 stations, and simulated maximum values were too low at a few stations for birch (but were still within several pollen grains m^{-3} of observed values). A more-detailed summary of the model evaluation is now given in Section 2.4, lines 593-605. We have added significant additional text to the revised manuscript (as described throughout this document) to assuage concerns about the appropriateness of the model approaches used and justify the specific choices made regarding which studies/models, forcing threshold values and base temperatures were assigned (e.g. Section 2.1, lines 160-185, 2.1.1, lines 188-190 and 234-239, pertaining to chilling species: Section 2.1.2., lines 300-305 and 338-374) We hope that in the revised manuscript it is clearer that we only applied literature-derived relationships between meteorological variables and flowering, as well as critical forcing and heating thresholds from previously published studies that are relevant to the species in question and representative of the climate zones within the domain (e.g. lines 104-107, 160-188, and 759-765). We also have placed more emphasis on the fact that the chosen thresholds, base temperatures, etc. are observationally constrained (lines 737, 760-763) which should serve to increase confidence in the model. Finally, we have placed greater emphasis on the fact that the relationships are valid only for simulations performed in similar climate zones (throughout the manuscript including changes to the abstract, lines 39 and 47; the introduction, lines 96-107; and the discussion/conclusion sections, lines 750-752 and 853-856.

"Third, I think that the chosen simulation period is not valid to support the interpretations drawn from the results. Even the authors admit that the differences in the potential pollen production between current and future climatic conditions could be due to a shift of the pollen season into or out of the simulation period. However, this is not further investigated. Hence, I recommend to use a more suitable simulation period."

R2 Author Response 3: We have re-run the future and current simulations for a full year and now present those results accordingly (this has been changed throughout the paper).

"Forth, the authors state that their model is flexible regarding different pollen species. But: for many species, they use identical formulations/fixed values with the hint that appropriate data for the given species is not available. Hence, the nature of the model might be flexible, but the use of identical parameters corrupts this flexibility. Especially data on birch seems to be scarce in Southern California since almost always data of other tree species is used for birch. I doubt very much that, e.g., for birches and olives the same parameters should be used since these trees primarily do not grow in the same climatic regions."

R2 Author response 4: We have endeavored to more fully describe that the STaMPS modeling framework is in fact flexible with respect to simulations in different climate zones because it includes several types of budburst models (thermal time, sequential, and alternating; Section 2.1, lines 160-185), and that in principal, GDD and chilling thresholds and base temperatures obtained from published studies of the selected species growing in other climatic regimes could be applied, and the choice of budburst model modified, in order to perform simulations in other locations (Section 4.2, lines 825-837). However it is also noted that for some species extensive additional phenological and aerobiological datasets would be required to inform and validate simulations performed in other regions (Section 4.2). STaMPS is designed such that modules can be modified and substituted not only for individual Plant Functional Types (PFTs), but also for species occurring in different climatic regions. This has been made clearer in Section 1 (lines 100-107), Section 2.1 (lines 174-185), and Section 4.2 (lines 825-837). Plans to incorporate additional climate-zone simulation capabilities as well as new species are discussed briefly in Section 4.2 (lines 822-840). Also as described in earlier author responses (above) we hope the changes made to the description of how and why the specific modeling approach was applied to birch and other chilling trees (Section 2.1.2, lines lines 300-305 and 338-374) justify inclusion of this species in the simulations.

"Overall, the paper leaves the impression that too many issues were tackled at the same time. The paper does not only present a phenological model describing beginning, end and course of a specific pollen season, but does so for several different taxa. Additionally, distribution maps for each of these taxa are generated. This paper could easily be divided into 2 or even more papers: one paper about the phenological model (or even one paper for each taxa including a thorough validation for each taxa) and one paper about the generation of distribution maps."

R2 Author Response 5: While we agree that given the breadth of the modeling effort, we knew that more than one manuscript would need to be prepared and submitted in order to publish a satisfactory description of the modeling framework, yet we could not identify a single journal that was appropriate to submit both the phenological model paper and the dispersal paper. Therefore we opted to submit to two EGU journals, and this has been quite a challenge in and of itself. However we also agree that the manuscript is too long and so we have moved the section describing determination of species composition, fractional vegetation cover, and land use in the domain (formerly Section 2.4) into an Appendix (A) so that the main body of the paper is more focused on the model itself. This seemed appropriate as this chapter is not at the heart of the model; i.e., other land- and vegetation-cover datasets could in principle be used to determine the necessary land cover input values needed for STaMPS simulations in a given gridded model domain.

"Introduction: I am missing a few sentences about existing models for the timing and magnitude of pollen seasons and their performance. The introduction presents the motivation for the study and a summary of the paper, but is missing a paragraph about the scientific context (with relevant reference)."

R2 Author Response 6: In lieu of including this in the Introduction, we have added the following text to the beginning of the section that discusses these approaches (Section 2.1, lines 160-185):

"A number of models have been devised to predict the timing of anthesis in trees and other PFTs that flower during spring to early summer (when temperature is the main driver controlling flowering); these include approaches that consider only the effects of forcing temperatures (such as the thermal time model; Cannell and Smith, 1983) and models that include both chilling and forcing temperatures: the sequential model (Sarvas, 1974), the parallel model (Landsberg, 1974), and the alternating model (Murray et al., 1989). Chuine at el. (1999) tested eight phenological models for their ability to predict flowering dates in tree species with differing abilities to adapt to local climatic conditions and found that no one model was best-suited for all species. STaMPS applies different models to different species based on whether the species has forcing-only or both forcing and chilling requirements for flowering. The thermal time model is applied to species without known chilling requirements (Section 2.1.1); a sequential model is applied to tree species with chilling requirements for flowering (Section 2.1.2). Section 2.2 describes an alternating model (based on Gleichsner and Appleby, 1996) that is applied to the *Bromus* grass species included in the initial model simulations. A particular model may perform well for simulating a species in one climate zone, but poorly for the same species growing in a different climate. The type of bud burst model assigned to a species in one location can be easily modified since STaMPS modules already exist for thermal time, sequential, and alternating approaches. STaMPS lacks a parallel budburst model although this may be ideal for predicting flowering in some species/locations. Linkosalo et al. (2008) found that sequential and parallel models had similar prediction accuracies for several tree species including two birch species in a high-latitude location (Finland), but that a simple thermal time approach performed best with independent data. Linkosalo et al. (2008) also noted that models including chilling parameterizations may be better suited to simulations under climate warming scenarios when chilling could potentially become a limiting factor."

"Page 2330, lines 1-2: what are the criteria for the selection of the species based on pollen count data?"

R2 Author Response 7: We have added the following text (lines 129-133):

"An aerobiological pollen dataset for 2003-2010 collected at the California Institute for Technology (CalTech) campus in Pasadena, CA was used for the selection of species to be included in the initial simulations along with the expertise of co-authors who have been studying local pollen in S. CA for a number of years. Only species with known allergenicities that typically flower in the March-June period in the study area were selected including species that occur naturally throughout the domain (such as oak) and those that occur mainly in urban environments in the domain but are also important allergens in other regions (e.g., birch)."

"Page 2330, lines 13-16: In some species (e.g. birch), the magnitude of pollen produced is not only a function of the meteorological conditions in the given season, but also depends on the previous season (the concept of masting). It is not clear whether this fact is included in the model or not."

R2 Author Response 8: We have deleted "in a given season" to avoid confusion since the point of this sentence was to list the variables known to govern pollen production but not the timescales on which they operate. In Section 2.3 of the manuscript we acknowledge that precipitation a full year prior to the start of pollen season in trees can affect the amount of pollen produced and that is in fact how STaMPS calculates the pollen production size for trees.

"Page 2331, lines 16-17: The authors exclude the year 2007 because of late-season rains. However, the section 2.1 addresses the prediction of the start of the pollen season. I don't see why late-season rains disqualify the year 2007 for the prediction of the start of the season which should be before the rainy late-season period."

R2 Author Response 9: Perhaps we were not clear in how we used the pollen count data and why 2007 was excluded: we have re-worded the explanation (lines 219-220) for the exclusion of this year from the set of years where cumulative GDD values were calculated and averaged by genus to determine GDD thresholds. Rains were so frequent in the early spring period that for many genera peaks were not observed at all during the expected times (or others). We note that a similar approach was taken in another study (Galán et al., 2001). In addition to hopefully making the exclusion and the reason for it more clear in the revised manuscript, the following text has also been added (lines 220-223):

"In evaluating several models for their ability to predict the start of *Olea* pollen season in Spain, Galán et al. (2001a) excluded two years out of an 18-year olive pollen dataset due to rain events occurring near the beginning of olive pollen season."

"Page 2331, lines 6-13, lines 23-24: Needs more justification how the specific base temperatures and start-dates for heat accumulation were chosen."

R2 Author Response 10: Please see author responses 2 and 6 above and 13 (below) which have addressed this.

"Page 2332, line 1: The GDD thresholds are taken as the average GDD values that were reached on the day of the peak pollen concentrations. However, the peak

pollen concentrations surely do not reflect the start of the pollen season. The start of the pollen season is a phenological date which depends on the meteorological conditions of previous days/months. What is the definition of the start of the pollen season used in the present study? The peak pollen concentrations depend on the current weather conditions and usually appear during the main season several days after the start of the pollen season. It should also be kept in mind that the start of the pollen season and the peak pollen concentrations could be a consequence of long-distance transport, not being related to the start of local flowering at all. Data should be corrected for these influences as far as possible."

R2 Author Response 11: We used the peak pollen concentration as date of peak flowering (as described in lines 216-239) and then distributed the season over a two-week period surrounding the peak date. We also acknowledge in the revised text the limitations of this approach. The following text was inserted (lines 258-277):

"A parameterization is applied to all simulated species that normally-distributes the pollen available for release on the calculated peak date of flowering over a two-week period (i.e., one week on either side of the calculated peak date) which was typical of the oak phenological observations and which collaborators in the present study have corroborated in their observations of temporal trends in pollen count intensities for a number of species, (not shown). Jato et al. (2007) observed flowering periods ranging from 8-13 days in populations of *Betula pendula* as well as in *B. alba* growing in Spain. Observations have indicated longer duration of flowering season in other species including sagebush (*Artemisia tridentata*) which has been observed to flower over a period of 4-5 weeks (Laursen et al., 2007) and bermudagrass (*Cynodon dactylon*), in which flowering has been observed to last 1-2 months depending on cultivar (Van De Wouw et al., 2009). There have also been reports of skewness in flowering distributions among populations (Laursen et al., 2007), as well as observations of temperature and precipitation effecting the shape of such distributions (Tedeschini et al., 2006). Parametrizations do not yet exist in STaMPS to represent environmental influences on flowering distributions, therefore the shape of simulated pollen curves may not be well-represented in years when flowering behavior within a population deviates from a normal distribution. On the other hand, STaMPS modules can be easily modified to represent longer flowering periods or different distributions to reflect averages observed for a given species."

With regards to the consideration of long-range pollen transport, we have inserted the following text into Section 4.1 (lines 739-765):

"Long-range transport of pollen may also complicate the interpretation of observed atmospheric pollen curves (e.g., Siljamo et al., 2008), especially for non-native species within a study domain that could have a much larger presence in regions outside of the domain, such as birch species. Jato et al. (2007) compared *Betula* pollen curves with phenological observations from two species of birch growing in Spain, and found that for the calculation of GDD requirements, using dates of peak pollen concentration yielded similar results as use of phenological observations, although lags between dates of peak flowering and peak pollen count have been reported for a number of species (e.g., Latorre, 1997). Although we cannot exclude the possibility that pollen transported over large ranges may have influenced the pollen count data used for the determination of heat thresholds, a multi-year aerobiological dataset such as that employed in the present study should be dominated by local signals over time, and numerous studies (of the same genera selected for initial simulation in STaMPS) performed in similar climatic regions were consulted during model development. In an effort to include the potential contributions from pollen transported into the domain during the evaluation simulation period, a nested domain scheme in which the 4-km S. CA domain pollen transport simulations used boundary conditions provided by simulation results of a larger 12-km domain that covers the entire states of California and Nevada (Figure 1; as described in Section 2.4) was used; the simulations results indicate contributions from California and Nevada outside the S. CA domain to pollen counts at the observational sites were insignificant (Zhang et al., 2013a). Our study is the first to predict pollen production for multiple species in the western half of the US, and is constrained with observations taken from within the study domain as well as built using numerous previously published relationships for the selected genera occurring in similar climate zones as the domain, therefore STaMPS improves capabilities for predicting pollen season in this region. Nonetheless, these limitations should be kept in mind when interpreting the pollen dispersal results.**"**

"Page 2332, lines 25-29, and page 2333, lines 1-2: It is mentioned that a 'variation-mimicking' parameterization has been developed and applied to several species. Details about this parameterization are not given. If I understand correctly, the parameterization normally distributes the available pollen over a period of 2 weeks. The authors state that such a behavior has been observed for a number of species, corresponding data is not shown. Looking at the pollen seasons in Europe, I cannot support the observation that typical pollen seasons have a length of only 2 weeks. For birch, e.g., the pollen season has been observed to be positively skewed instead of normally distributed (see e.g., Grewling et al. 2012, Grana). Additionally, it should be taken into account that the length of the pollen season is influenced by the weather conditions, e.g. cold conditions tend to result in a longer than usual pollen season. As I understand from the paper, this is not taken into account."

R2 Author Response 12: See author response 11, above.

"Page 2333, lines 3-15: Instead of using TB values of another species, a species-specific TB value for the studied region could be found by systematically varying the base temperatures and the starting date of the accumulation. Otherwise, is there a justification why the values of another species are valid for the given species?"

R2 Author response 13: we have addressed this above and now the justifications are provided for the choice of base temperatures. With respect to systematically varying the base temperature we did experiment with this as is now described as follows (lines 348-374):

"T_B values close to 0°C have been suggested for birch growing in high latitudes although latitudinal gradients in base temperatures as well as chilling requirements have been observed in many tree species including birch (Myking and Heide, 1994) and grasses (Heide, 1994). We tested a range of base temperatures for birch using the Pasadena data but found little difference in terms of percent standard deviation in accumulated GDD on peak birch pollen count dates using a base temperature of 0°C compared with 9.1 for the years included in the analysis. Most studies of *Betula* have been performed on high-latitude ecotypes although Jato et al. (2007) studied and modeled the onset and length of flowering in *B. pendula* and *B. alba* populations in Spain using several modeling approaches and base temperatures and found that model parameterizations developed for olives (Galán et al., 2001b) resulted in the lowest deviations from actual versus predicted peak pollen date for *Betula*. Data regarding optimal T_B values for birch and walnuts growing in climate zones similar to the study domain are sparse therefore, following a similar approach as Jato et al. (2007), we applied a model developed for olives (De Melo-Abreu et al., 2004) to birch and walnut species within the domain. It should be noted that Jato et al. (2007) used the olive model of Galán et al. (2001b), while we have employed the De Melo Abreu et al. (2004) approach which allows for the calculation of devernalization during chilling calculations. De Melo Abreu et al. (2004) evaluated the inclusion of devernalization against the same chilling calculation approach used by Jato et al (2007) and found results to be more accurate, and the model more physiologically meaningful, when devernalization was considered. It has been long known that species with chilling requirements can lose a portion of their accumulated chilling when temperatures exceed some threshold value (Richardson et al., 1974; Gilreath and Buchanan, 1981). The GDD thresholds for these species were determined using the Pasadena pollen counts as described in Section 2.1.1, with GDD accumulation dates each year beginning on the date when chilling requirements were met for each chilling species according to the Pasadena meteorological data."

"Page 2334, lines 22-25: Any justification why you use the olive value for birch instead of the walnut value?"

R2 Author response 14: This is addressed in the author responses 2, 6, and 13, above.

"Page 2335, lines 3-6: Maxima/minima in pollen concentrations can have different reasons: e.g., precipitation events washing out the pollen in the air, turbulence and wind strengthening/weakening the emission and diffusion of pollen, long-distance transport, variations between individual trees. I find it implausible to assign observed maxima in airborne pollen concentrations mainly to different olive cultivars and individual thermal requirements."

R2 Author response 15: We have acknowledged that multiple factors could have contributed to the observed phenomenon and we have further explained why we think multiple olive cultivars are the reason by adding the following text (lines 376-392):

"Several maxima are often observed in the Pasadena olive pollen data in a given year (Fig. 2) and these do not appear to be correlated to precipitation events (not shown). Some explanations for this phenomenon include (1) the effects of changing wind fields causing times of maximum pollen release to be missed or obscured, (2) the possibility of long-range pollen transport (discussed in Section 4.1), or (3) the presence of several olive varieties with different thermal requirements for flowering. As mentioned previously, the Pasadena pollen data was collected at CalTech; this campus has a large and diverse collection of olive trees growing on its grounds, some of which are used to make oil. The proximity of this known source of *Olea* pollen near the sampling site makes it likely that local effects were dominating observed pollen concentrations. De Melo-Abreu et al. (2004) found that different olive cultivars have unique chilling and heating requirements, but since the datasets used to determine tree species composition for this study (described in Appendix A) identified all olives present within the domain as simply "*Olea europaea*", we applied the average heating and cooling threshold values across all of the cultivars studied by De Melo-Abreu et al. (2004), which yielded the same threshold value that was calculated for olives using the Pasadena data (Table 1)."

"Page 2337, lines 2-6: Relevance of these remarks? Should be part of the section 'Future plans'."

R2 Author response 16: We have deleted this text. (These plans are discussed in Section 4.2)

"Page 2338, lines 16-19: Is it wise to use model precipitation to construct the relationship between precipitation and the pollen potential? It is well known that precipitation is one of the parameters that are usually not very well simulated in models. I suggest to use observations for that purpose."

R2 Author response 17: The PRISM model ingests numerous observational data from stations all over the U.S. including many in the western half of the country, and as such is well-constrained. I would agree that maybe using a 1-year precipitation map for PRISM would be risky but we used a 30-year average dataset, and this product is, in the opinion of the authors, one of the best precipitation datasets available.

"Page 2338, line 22: Which other trees? Is the function also based on data of these other trees or is it taken from oak data? If taken from oak data: justification?"

R2 Author response 18: We have changed the text to (lines 498-500):

"The relationship between precipitation and pollen potential developed for *Quercus* is also applied to *Morus*, *Platanus*, and *Juglans* and is expressed as…"

and (lines 505-510): "This approach may not accurately represent the represent the relationship between pollen production and prior-year precipitation for the non-oak species that it was applied to (as observed values were not available for comparison for these other species), but probably broadly represents the positive value between pollen production and prior-year precipitation observed for numerous tree species (Fairley and Batchelder, 1986; Kozlowski, 1971)."

"Page 2339, lines 22-25: Justification for neglecting the influence of precipitation on the length of the flowering season? Although the simulated grass species were not included in the mentioned paper, it is very plausible that the discovered influence also plays a role here."

R2 Author response 19: We have added some discussion of this and a consideration of limitations associated with this assumption (lines 269-277):

"There have also been reports of skewness in flowering distributions among populations (Laursen et al., 2007), as well as observations of temperature and precipitation effecting the shape of such distributions (Tedeschini et al., 2006). Parametrizations do not yet exist in STaMPS to represent environmental influences on flowering distributions, therefore the shape of simulated pollen curves may not be well-represented in years when flowering behavior within a population deviates from a normal distribution. On the other hand, STaMPS modules can be easily modified to represent longer flowering periods or different distributions to reflect averages observed for a given species."

"Page 2340, line 10: Please add the p-value of the correlation."

R2 Author response 20: The p-value has now been added to the caption of Fig. 5 (instead of adding it to the text).

"Page 2340, lines 12-14: How did you calculate this?"

R2 Author response 21: The text has been changed to read (lines 557-563):

"When percent early-season vernalization (as defined below) was plotted against observed peak pollen concentrations (normalized as percent of average peak value across all years), an r^2 value of 0.75 was observed using a linear least-squares regression, while the same regression performed between prior-year wet season precipitation and observed peak pollen concentrations (normalized as above) yielded an r^2 value of only 0.08 (not shown)."

"Section 2.4: I am missing the details: how is the fractional land cover assigned (percentages used for each class!), what is used for the weighting, horizontal resolution of the data sets: overall, I do not understand how the fractional land cover was produced. Maybe a flow chart would clarify the process? It would also be good to compare the methods with the literature about vegetation cover estimation (e.g., Sofiev et al. 2006, Int. J. Biometeorology, or Skjoth et al. 2008, Ecological Modeling, or Skjoth et al. 2010, Agricultural and Forest Meteorology, or Pauling et al. 2012, Int. J. Biometeorology)."

R2 Author response 22: We have removed the word "weighted" in describing how the urban tree inventories were used to calculate species composition since these were simply averaged, not weighted. We have provided spatial resolution for the CDL raster dataset used to determine land-use and (for agricultural locations) species composition. We have also provided spatial resolution for the NLCD percent canopy cover and impervious raster datasets used to determine fractional vegetation cover. As already described in this section (which is now an appendix), the FIA dataset was a point shapefile while the EPA ecoregions dataset and the NRCS dataset were both polygons. To clarify the process of assigning fractional species composition within domain cells, we have added the following paragraph (in lieu of adding another figure to the paper):

"For example, suppose a given domain cell is comprised of 20% low-density urban land use, 70% forest, and 10% water according to the CDL dataset. The urban portion of the cell would be assigned a static 15% canopy cover (Table 4) and the associated fractional species composition for that area would be assigned based on an urban tree inventory. The forested portion would be assigned the fractional tree cover obtained by averaging the NLCD canopy dataset underlying the forested portion of the cell, while species composition for the forested region would be assigned based on which ecoregion(s) occur within the cell and weighted accordingly. Tree species composition within different ecoregions was determined by averaging the FIA data described above according to ecoregion membership across the FIA plots. The portion of the cell identified as being covered by water would be excluded from analysis."

We feel that the extensive revisions have already lengthened our already long paper considerably; therefore we opted not to compare our fractional species cover determination approach with other studies. Additionally we have moved this section into an appendix in order to focus the main body of the paper on the pollen model itself and not on the land cover datasets employed; in principal various other datasets or techniques could be applied to determine species composition.

"Page 2342, lines 18-23: This is not necessary here, already described in formula 3."

R2 Author response 23: We have deleted this text.

“Section 2.5: Before applying the model on future climate, I would expect some sort of validation using current climate. Page 2344, lines 10-17: How well works the model for current climate? In order to interpret results for future climate, it is essential to know how precise the timing and magnitude of the pollen season can be calculated under current climate.”

R2 Author response 24: See Author responses 2 and 11 (above) and 26 (below).

“Page 2345, lines 20-end of paragraph: Maybe the simulation period should be extended to represent the entire pollen season? It would help to interpret the results if the entire season was inside the simulation period!”

R2 Author response 25: We have rerun the current & future simulations for a full year period and have modified the manuscript accordingly

“Page 2346, lines 5-14: To evaluate the model, STaMPS output was incorporated into a transport model and simulated pollen concentrations were compared to pollen count data. The reader is referred to a companion paper for the details. However, I think incorporating the model output into another model and comparing the output of the second model to observational data is not a good way to validate the first model. Resulting pollen concentrations do not only depend on the output of the STaMPS model but also on the emission parameterization and transport/diffusion processes in the transport model. If available, it would be nice to compare the STaMPS output directly to phenological data of the start of flowering.”

R2 Author response 26: We agree that phenological observations are the ideal datasets to inform and validate predictive models of budburst, and many of the studies that we used to parameterize these relationships for the selected species were based on phenological observations (e.g. *Bromus*: Gleichsner, J. A. and Appleby, A. P. Weed Sci., 44, 57-62, 1996; *Quercus*: phenological observations provided courtesy of Dr. Walt Koenig of Cornell University; and Olea: De Melo-Abreu, J. P., Barranco, D., Cordeiro, A. M., Tous, J., Rogado, B. M. and Villalobos, F. J. Agr. Forest Meteorol., 125, 117-127, 2004). As mentioned in an author response, we have now included a discussion of the possible contribution to observed pollen from long-range transport (LRT) episodes and the steps we took to consider LRT in our simulations as well as acknowledging the limitations inherent in using observed pollen counts as proxies for dates of peak pollen release (Section 4.1). Finally, as mentioned previously, while we agree that additional model evaluation is always desirable, our companion paper did evaluate the results rather rigorously despite the use of pollen counts instead of phenological observations. As pointed out in our companion paper (Zhang et al., 2013), although LRT might be expected to affect “spatial distribution and magnitude of pollen concentrations at the local scale (Zink et al., 2011), there should be a correlation between the timing of pollen emission and that of concentration on the regional scale”, and using birch performance as an example, the regionally-observed and simulated regionally-averaged mean peak pollen concentration dates both occurred in Mid-May, and comparisons between observed mean and max concentrations for nine sites within the domain indicated good performance for most of the species (Also see Author responses 2 and 11, above).

"Page 2347, lines 18-24: I totally agree! Please justify why you did not simulate the entire season. Regarding the limitations (not representing the entire pollen seasons), I am not sure what we can learn from the study about the future."

R2 Author response 27: We have rerun the current & future simulations for a full year period and have modified the manuscript accordingly

"Page 2349, lines 8-9: No, the STaMPS model has been designed to simulate the timing and potential magnitude of the pollen season. The release of the pollen is simulated in the transport model (see companion paper) using an emission parameterization respecting the influence of wind."

R2 Author response 28: We have changed "release" to "production" (thanks for catching that)

"Figure 3: What is "Days to flower" (y-axis)? The beginning of the pollen season? The length of the pollen season? If it is the beginning: what is the initial date? When does the counting of the days (x-axis) start? What is 7°C? Mean/min/max temperature?"

R2 Author response 29: We have added the following sentence to the caption of Fig. 3:

"*B. diandrus s*eeds were exposed to cold treatments and germinated during the chilling period; subsequent days to flowering following vernalization treatments are shown on the y-axis."

"Figure 4 and 5: What is the p-value of the correlation?"

R2 Author response 30: We have added p-values to the captions in Figs. 4 and 5.

"Technical corrections:"

"Page 2333, lines 20 + 23: (Betula, Juglans, and Olea) : : :(olive, walnut, and birch):

unnecessary repetition, additionally it would be better to use either Latin or English words, not a mixture."

R2 Author response: we have deleted "(olive, walnut, and birch)"

"Page 2334, line 7: Losing a negative value (-0.56 chilling units) results in a net gain: double negative = positive: : :in my opinion, it should be: : : :above which 0.56 chilling units are lost."

R2 Author response: you are right; this has been fixed

"Page 2335, lines 24-30: illogical use of the words 'quantitative' and 'qualitative': for me, the fraction of heads flowering is quantitative as it gives the amount of potentially available pollen grains."

R2 Author response: we have changed this sentence to read:

"In grasses, vernalization may be *facultative*, where cold temperatures occurring prior to the flowering season affect the timing of flowering and can enhance the fraction of plants that head, or *obligate*, in which adequate chilling is required for flowering to occur at all (Gleichsner and Appleby, 1996)."

"Page 2339, formula 5: the precipitation-driven coefficient is now called gamma. Before, in formula 3, it was called alpha. There, gamma was the fraction of land covered with the specific plant."

R2 Author response: We have changed this to alpha.

"Page 2349, line 22: Artemisia instead of Artemesia."

R2 Author response: We have changed this to Artemisia

"Table 1: phenological instead of phonological"

R2 Author response: Thanks for catching this; it has been changed.

Author response to Anonymous Referee #3, 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. (Geosci. Model Dev. Discuss., 6, C1029–C1032, 2013)

"In the manuscript there are appropriate references to a companion methodological paper, but the difference between pollen production at plant level and airborne pollen counts is not well explained here and this is necessary. A reader non expert in the field could assume, reading the current manuscript, that the pollen present in the air is simply the result of the pollen released by local plants, and this assumption is obviously wrong. Simulated pollen production data are in fact used as "input" data to be used by a transport model presented in the companion paper. I suggest to make a clear distinction between these two research aspects also in this ms. The confusion on this argument is also generated by the presence of Fig. 2, Fig. 4 and Fig.5, which report pollen counts while captions refer to pollen production, and the claimed use of pollen counts to calibrate a coefficient of the presented pollen production model (p.2338, lines 13-16). Such calibration also needs then a more accurate justification."

R3 Author response #1: We have changed the beginning of Section 2.3 (lines 448-452) to read:

The amount of pollen emitted to the atmosphere is a function both of the amount of pollen available for release (which is what is modeled by STaMPS) and the various short-term meteorological factors that influence pollen dispersal into the atmosphere, including relative humidity, wind characteristics, etc. (as simulated by the companion pollen transport model described in Zhang et al., 2013a).

In response to the issue with Figure 2, we cannot find a discrepancy between the figure caption and the figure itself. However, we have made the caption more explicit by changing it from 'Olive tree pollen counts data for Pasadena, CA, 2006 and 2008.' to 'Olive tree pollen concentrations observed in Pasadena, CA, 2006 and 2008.' In Figure 4, we have changed the caption to read: "Relationship between wet season precipitation and pollen concentrations derived from Fairley and Batchelder (1986, who reported a p-value of 0.0002). This relationship is applied to pollen production by tree species in the domain without chilling requirements." We have changed the caption of Figure 5 to read: "Relationship between pollen production (expressed as percent of average peak value) and dual vernalization-precipitation coefficient (as calculated using Eq. 6) derived from Pasadena, CA pollen concentration data and applied to STaMPS tree species with vernalization requirements for flowering."

Also see lines 376-392 where we have made it clearer to the reader that observed pollen concentrations can be affected by various meteorological factors. We have inserted the following text into Section 4.1 (lines 739-765):

"Long-range transport of pollen may also complicate the interpretation of observed atmospheric pollen curves (e.g., Siljamo et al., 2008), especially for non-native species within a study domain that could have a much larger presence in regions outside of the domain, such as birch species. Jato et al. (2007) compared *Betula* pollen curves with phenological observations from two species of birch growing in Spain, and found that for the calculation of GDD requirements, using dates of peak pollen concentration yielded similar results as use of phenological observations, although lags between dates of peak flowering and peak pollen count have been reported for a number of species (e.g., Latorre, 1997). Although we cannot exclude the possibility that pollen transported over large ranges may have influenced the pollen count data used for the determination of heat thresholds, a multi-year aerobiological dataset such as that employed in the present study should be dominated by local signals over time, and numerous studies (of the same genera selected for initial simulation in STaMPS) performed in similar climatic regions were consulted during model development. In an effort to include the potential contributions from pollen transported into the domain during the evaluation simulation period, a nested domain scheme in which the 4-km S. CA domain pollen transport simulations used boundary conditions provided by simulation results of a larger 12-km domain that covers the entire states of California and Nevada (Figure 1; as described in Section 2.4) was used; the simulations results indicate contributions from California and Nevada outside the S. CA domain to pollen counts at the observational sites were insignificant (Zhang et al., 2013a). Our study is the first to predict pollen production for multiple species in the western half of the US, and is constrained with observations taken from within the study domain as well as built using numerous previously published relationships for the selected genera occurring in similar climate zones as the domain, therefore STaMPS improves capabilities for predicting pollen season in this region. Nonetheless, these limitations should be kept in mind when interpreting the pollen dispersal results."

As described in Section 2.3 of the manuscript, the pollen production potentials were derived from studies of actual pollen production; i.e. production of pollen grains was quantified and scaled to a per-unit area production potential, which is then modified by a precipitation-only or a combined precipitation/vernalization coefficient. Admittedly the coefficients were derived from pollen concentration data, in our case peak concentrations were used, and the relationship between these and early-season chilling/precipitation were applied to the potential pool size of pollen produced. While we agree this approach leaps from production to concentration, the observation in aerobiological datasets that cumulative past-year meteorological conditions such as precipitation affect pollen concentrations is pervasive and overwhelmingly argues that these factors control the amount of pollen available for release while on shorter timescales local meteorological factors contribute to pollen rafting into the atmosphere (e.g. Laursen et al., 2007 from the manuscript). In Fairley and Batchelder (1986), the study whose observations of the relationship between pollen concentrations and wet-season precipitation we used to derive our precipitation coefficient, an excellent correlation was observed between median pollen concentration and wet-season precipitation a full year before the year of flowering, while median pollen concentration and wet-season precipitation in the year of flowering had no statistically-significant relationship. This behavior illustrates that meteorological factors occurring well before the date of flowering (i.e., during the times when flower bud differentiation is occurring) are controlling these processes, which has also been reported in phenological studies numerous times (Corden and Millington, 1999; Aerobiologia, 15; Miyazaki et al., 2009 from the manuscript; Kozlowski, 1971 from the manuscript), and which justifies the application of observed aerobiological relationships with prior-year climatic variables to predicted production. We have made it more explicit that pollen concentration data were used to infer the relationships between meteorological factors and pollen production (see changes made above) in tree species (the grass pollen production coefficient was determined based on phenological data) and have also endeavored to more clearly discern between what STaMPS does (i.e., predict the amount of pollen available to be released) and what the pollen transport model described in our companion paper does (model the dispersal into the air).

"The manuscript refers to the interaction between airborne pollen and anthropogenic air pollutants such as for example ozone. But no information about the ozone cycle is given, even if the concomitance of peak ozone concentration and

pollen concentration is mentioned. I would better describe this point, to highlight the occurrence at the same time of different airborne particles and justify the choice of the period March-June for simulations."

R3 Author response #2: As described in lines 584-589 "Initial STaMPS and pollen transport simulations were performed for a 4-km resolution model domain centered over southern California (Figure 1, lower right corner) for 1 March through 30 June 2010. This domain and time period coincided with an extensive set of pollen observations collected as part of the University of Southern California's Children's Health Study from which ambient pollen count data were collected and were used to evaluate and optimize STaMPS."

Although ozone season frequently occurs in or near the selected simulation period in this region, we felt it would be inappropriate to add a discussion of this in light of the fact that this is a complex issue (i.e. ozone episodes can occur at various times depending on numerous factors) and also because the extensive revisions that have been made to the paper (see Author responses to Referee # 1 and #2) have considerably lengthened an already long manuscript, additionally this really isn't the focus of the manuscript. One of the other reviewers even thought the paper could be submitted as two papers but as we discussed in response to that reviewer we didn't feel this was an option for our case (mainly due to the existence of our companion paper which was submitted also to an EGU journal and which is under review now).

"I am very perplex about the use of the same chilling requirement for olive and birch. The authors have probably misunderstood the papers cited to justify such an arbitrary choice. There are several points that can be mentioned in order to consider such an attribution wrong, among them a very different latitudinal optimum between these two genera (birch is more acquainted to higher latitudinal range -Northern Europe- than that considered in the presented ms, differently from olive for which the considered geographical extent could be optimal), as well as a different ecological classification (olive is a late successional tree, birch an early successional one). I would reconsider this aspect in order to modify the model."

R3 Author response #3: We have now explained that specific model approaches for birch (growing in the climates within our study domain, i.e., far from their latitude of origin) are chosen because these have proven to be the best for simulation of birch behavior in similar climates, by inserting the following text into the manuscript (lines 300-305)"

"Of the tree genera selected for simulation with chilling requirements for flowering (*Betula*, *Juglans*, and *Olea*), olives (*Olea europaea*) have been best studied in Mediterranean climate zones since they are native to these regions and are important not only economically but are also highly allergenic (Rodríguez et al., 2001). The sequential chill-heating model developed for olives by De Melo-Abreu et al. (2004) is used to simulate pollen season for tree species in STaMPS with chilling requirements, since sequential models have been identified as being appropriate for phenological simulations of both birch and olive in similar climates (Jato et al., 2007; De Melo Abreu, 2004)."

and (lines 338-374) "For olive, walnut and birch species, we use the same optimum and breakpoint chilling temperature values as those selected for olives in De Melo-Abreu et al. (2004), since these values were determined within a similar climate zone as the present study, and observations (e.g., Jato et al., 2007; Warmund et al., 2009) suggest that birch and walnut trees have similar optimal chilling temperatures as olives. A threshold chilling quantity of 58 chilling units was selected for walnuts (Warmund et al., 2009); 432 chilling units was assigned to olives (De Melo-Abreu et al., 2004) and was also assigned to birch (for which published values for birch growing in similar climates were not available).

T_B values close to 0°C have been suggested for birch growing in high latitudes although latitudinal gradients in base temperatures as well as chilling requirements have been observed in many tree species including birch (Myking and Heide, 1994) and grasses (Heide, 1994). We tested a range of base temperatures for birch using the Pasadena data but found little difference in terms of percent standard deviation in accumulated GDD on peak birch pollen count dates using a base temperature of 0°C compared with 9.1 for the years included in the analysis. Most studies of *Betula* have been performed on high-latitude ecotypes although Jato et al. (2007) studied and modeled the onset and length of flowering in *B. pendula* and *B. alba* populations in Spain using several modeling approaches and base temperatures and found that model parameterizations developed for olives (Galán et al., 2001b) resulted in the lowest deviations from actual versus predicted peak pollen date for *Betula*. Data regarding optimal T_B values for birch and walnuts growing in climate zones similar to the study domain are sparse therefore, following a similar approach as Jato et al. (2007), we applied a model developed for olives (De Melo-Abreu et al., 2004) to birch and walnut species within the domain. It should be noted that Jato et al. (2007) used the olive model of Galán et al. (2001b), while we have employed the De Melo Abreu et al. (2004) approach which allows for the calculation of devernalization during chilling calculations. De Melo Abreu et al. (2004) evaluated the inclusion of devernalization against the same chilling calculation approach used by Jato et al (2007) and found results to be more accurate, and the model more physiologically meaningful, when devernalization was considered. It has been long known that species with chilling requirements can lose a portion of their accumulated chilling when temperatures exceed some threshold value (Richardson et al., 1974; Gilreath and Buchanan, 1981)."

Additional specific remarks:

"p. 2327, line 2: change "A pollen model..." to "A pollen production model...", or "A model of pollen shedding/production...", or simply "A model..."."

R3 Author response: We have changed this to "A model…"

"line 3: Is there need to specify "terrestrial"?"

R3 Author response: We have deleted "terrestrial"

"line 4: The model does not strictly study the interaction between pollutant factors and pollen, but simply simulates airborne pollen distribution. I would rather suggest "in order to investigate how pollen can interact with anthropogenic pollutants to affect human health". "

R3 Author response: We have changed this to "in order to investigate how pollen can interact with anthropogenic pollutants to affect human health"

"line 12: Are references to "Mediterranean zones" in Southern California appropriate? Would not be better a more general climatic definition? The adjective "Mediterranean" also recurs in other parts of the manuscript. When not referring to plant species, it would rather be better to use a different expression."

R3 Author response: This region is frequently referred to as being Mediterranean climate-wise; in fact much of the state of California is included in maps of Mediterranean climates. We have, however, removed numerous instances of this wording and simply replaced with phrases such as "from studies performed in similar climates", etc.

"p. 2328, line 3: change "simulated species" to "considered species"."

R3 Author response: We have changed this to "considered species"

"line 25-26: please also provide more recent literature about advances in phenology timing, for example Menzel et al. 2006, and about a detected changes in airborne pollen burden, for ex-ample Ziello et al. 2012, Damialis et al. 2007, Garcia-Mozo et al. 2010."

R3 Author response: We have made several additional references and comparisons to both other model results and observations (including Ziello et al., 2012 and others), see lines 643-646, 660-670 and 687-690 of the revised manuscript.

"p. 2330, line 20: Not clear, maybe "separate" stands for "separated"? If so, in which ways are they separated? Are those modules sequential? Parallel?"

R3 Author response: "separate" is correct; but we have now inserted the following text at the beginning of Section 2 (lines 160-185):

"A number of models have been devised to predict the timing of anthesis in trees and other PFTs that flower during spring to early summer (when temperature is the main driver controlling flowering); these include approaches that consider only the effects of forcing temperatures (such as the thermal time model; Cannell and Smith, 1983) and models that include both chilling and forcing temperatures: the sequential model (Sarvas, 1974), the parallel model (Landsberg, 1974), and the alternating model (Murray et al., 1989). Chuine at el. (1999) tested eight phenological models for their ability to predict flowering dates in tree species with differing abilities to adapt to local climatic conditions and found that no one model was best-suited for all species. STaMPS applies different models to different species based on whether the species has forcing-only or both forcing and chilling requirements for flowering. The thermal time model is applied to species without known chilling requirements (Section 2.1.1); a sequential model is applied to tree species with chilling requirements for flowering (Section 2.1.2). Section 2.2 describes an alternating model (based on Gleichsner and Appleby, 1996) that is applied to the *Bromus* grass species included in the initial model simulations. A particular model may perform well for simulating a species in one climate zone, but poorly for the same species growing in a different climate. The type of bud burst model assigned to a species in one location can be easily modified since STaMPS modules already exist for thermal time, sequential, and alternating approaches. STaMPS lacks a parallel budburst model although this may be ideal for predicting flowering in some species/locations. Linkosalo et al. (2008) found that sequential and parallel models had similar prediction accuracies for several tree species including two birch species in a high-latitude location (Finland), but that a simple thermal time approach performed best with independent data. Linkosalo et al. (2008) also noted that models including chilling parameterizations may be better suited to simulations under climate warming scenarios when chilling could potentially become a limiting factor."

"line 26: Please provide some literature to support the statement that "temperature is the main driver controlling flowering", for example Parmesan et al. 2007 or Menzel et al 2006."

R3 Author response: We already cited two studies the first time we stated that temperature was the main driver controlling flowering (lines 144-145 of the revised manuscript).

“p. 2349, line 22: change "Artemesia" to "Artemisia".”

R3 Author response: Thanks for catching this; it has been corrected

“Table 1, footnote: change "phonological" to "phenological".”

R3 Author response: Thanks for catching this; it has been corrected