Response to Referee #1

Thank you for your compliments and constructive suggestions to improve the manuscript. Referee #1's comments are individually listed below with a corresponding author response. Line numbers in author responses correspond to the revised manuscript.

Comment: General: Throughout the manuscript the word "climatological" or "climatology" is used instead of the word "average" or "mean" with regard to pollen. For example, lines 161- 162 state that: "For deciduous broadleaf forest (DBF) taxa, the Southeast has the highest climatological pollen maximum reaching up to about 700-1200 grains m-3 around day 100." This is confusing because it is applied to a non climatic/meteorological variable, and because it is used in a manuscript which also focusses on climate. It would be much better to simply state: "For deciduous broadleaf forest (DBF) taxa, the Southeast has the highest average pollen maximum reaching up to about 700-1200 grains m-3 around day 100."
Similarly, lines 159-160 could be modified from: "Figure 2 shows the observed clima- tological PFT daily pollen counts averaged over all stations within the defined subre- gions." to: "Figure 2 shows the observed average daily PFT pollen counts averaged over all stations within the defined subregions.".
And so on.

Response: All instances of "climatological" or "climatology" with regards to pollen emission fluxes and counts have been changed to "average", sometimes with added specificity (e.g, "8-year average pollen time series", Line 164).

Comment: Line 17, Abstract: "PFT" is used without being given in full earlier in the Abstract, so please provide both the full and abbreviated form here.

Response: PFT defined as "plant functional type" at first use in the in abstract (Line 17).

Comment: Line 40, Introduction: The authors may wish to refer to two recently published works that relate to the introductory material here and/or elsewhere in the Introduction:
Sofiev M, Prank M. Impacts of climate change on aeroallergen dispersion, transport, and deposition. In: Beggs PJ (Editor). Impacts of Climate Change on Allergens and Allergic Diseases. Cambridge University Press, Cambridge, 2016. pp 50-73.
Beggs PJ, Šikoparija B, Smith M. Aerobiology in the International Journal of Biometeorology, 1957–2017. International Journal of Biometeorology 2017. DOI: 10.1007/s00484-017-1374-5 [see section on "Aerobiological modelling and forecasting"]

Response: Beggs et al. 2017 is now cited on a new line that reads, "The interest and growing wealth of knowledge of allergenic pollen is reviewed by *Beggs et al.* (2017)." (Lines 40-41).
Sofiev & Prank 2016 was found useful for the discussion of climate-scale pollen dispersion models, thus it cited in a new line of the introduction, "Only recently have regional-scale modeling studies of pollen dispersion been conducted for Europe, and they have been used to assess the impacts of climate change on airborne pollen distributions." (Lines 45-47).

Comment: Line 98: This sentence makes reference to "the Finnish emergency modeling system (SILAM)". It would be better to change this to "the Finnish System for Integrated modeLling of Atmospheric coMposition (SILAM)" as given in the Introduction section of Sofiev et al. (2013).

Response: This correction to the acronym "SILAM" has been made on Lines 110-111.

Comment: Line 130, paragraph 1 of section 2.1: Table 1 does not relate to "NAB pollen count data ranging from 2003-2010 at all stations in the continental United States", so delete reference to it and just refer to Figure 1.

Response: The reference to Table 1 was intended to be a reference to Table S1; Table S1 is now referenced at the end of paragraph 1 of section 2.1, and any reference to Table 1 has been deleted.

Comment: Line 139: This line includes a reference to Table 2. There is no Table 2 in the manuscript. Should it be Table S2?

Response: The reference to Table 2 was intended to be toward Table S2. It has been corrected on Line 152.

Comment: Line 140: Change "Cupresseceae" to "Cupressaceae".

Response: This spelling correction was made on Line 150.

Comment: Lines 155-159, paragraph 1 of section 2.2: Currently just two of the four boundaries of each of the five subregions are provided. Please provide upper and lower limits of both latitude and longitude for each subregion.

Response: We have revised the subregion boundaries to include full boundaries as displayed in Figure 1. (Lines 167-173).

Comment: Lines 161-170: The paragraph in these lines seems to contain several values that do not match what is shown in Figure 2. Specifically, deciduous broadleaf forest (DBF) taxa in the Southeast does not have an average pollen maximum reaching up to about 700-1200 grains m-3. Figure 2b shows that it only reaches up to about 500 grains m-3. In the Northeast, DBF does not reach up to an average of 400 grains m-3. It peaks just above 240 grains m-3. And finally, a sharp maximum of 775 grains m-3 does not appear in the Mountain subregion. The sharp maximum is only about 360 grains m-3. The paragraph should be carefully checked.
Line 173: As above, please check the numbers 400 and 200 in this line.

Response: We thank the reviewer for drawing this to our attention. Figure 2 data was found to have errors, and as a result does not match the text as noted by the reviewer. We have revised Figure 2 with the correct data and now the figure is consistent with the manuscript text.

Comment: Lines 184-186: The discussion regarding C3 and C4 grasses here and/or elsewhere in the manuscript may be enhanced through reference to the following article:
Medek DE, Beggs PJ, Erbas B, Jaggard AK, Campbell BC, Vicendese D, Johnston FH, Godwin I, Huete AR, Green BJ, Burton PK, Bowman DMJS, Newnham RM, Kate- laris CH, Haberle SG, Newbigin E, Davies JM. Regional and seasonal variation in airborne grass pollen levels between cities of Australia and New Zealand. Aerobiologia 2016;32(2):289-302. DOI: 10.1007/s10453-015-9399-x

Response: Thank you for the suggestion. Medek et al. 2016 has been cited in the main discussion of C_3 and C_4 grass phenology in Section 2.2. New lines of text are as follows: "Similarly, *Medek et al.* (2016) observed two grass pollen peaks in Australia, with a stronger, late-summer peak at lower Southern latitudes where there is higher incidence of C_4 grass. However, the authors note that sometimes this may be due to a second flowering of some C_3 grass species." (Lines 211-213). This paper was also used in the discussion of phenological trends in Section 4.2 with new text as follow: "Trends for grass in Australasia show that the correlation of the end date of the pollen season with average spring temperature is positive, while the same relationship for the start date is negative, suggesting also that season start dates are earlier and season duration increases with warmer climates (Medek et al. 2016)." (Lines 408-410).

Comment: Line 278: Change "met" to "meteorological".

Response: Correction made from "met" to "meteorological". (Line 339).

Comment: Line 343: With respect to the Parry et al. 2007 citation, it should be Confalonieri et al. 2007 because the former are the book editors and the latter are the chapter authors. Change also in the References list (i.e., chapter authors first, book editors later in the reference). Further, here and/or elsewhere in this paragraph (lines 342-352) could benefit from reference to the following:
Ziska LH. Impacts of climate change on allergen seasonality. In: Beggs PJ (Editor). Impacts of Climate Change on Allergens and Allergic Diseases. Cambridge University Press, Cambridge, 2016. pp 92-112.
Ziska L, Knowlton K, Rogers C, Dalan D, Tierney N, Elder MA, Filley W, Shropshire J, Ford LB, Hedberg C, Fleetwood P, Hovanky KT, Kavanaugh T, Fulford G, Vrtis RF, Patz JA, Portnoy J, Coates F, Bielory L, Frenz D. Recent warming by latitude associated with increased length of ragweed pollen season in central North America. Proceedings of the National Academy of Sciences of the United States of America 2011;108(10):4248– 4251. DOI: 10.1073/pnas.1014107108

Response: The reference to Parry et al. 2007 has been changed to Confalonieri et al. 2007 as suggested. The additional references are excellent examples of general and advanced discussions on pollen phenology. Ziska 2016 is cited several times throughout the manuscript : **"**The spatiotemporal heterogeneity of climate change may affect which regions and seasons will be most influenced by climate change (Ziska 2016)." (Line 415-416), with a second citation on Line 419-420 and a third citation in the introduction: "Climatic changes in large-scale pollen distributions are mostly absent from scientific literature, though multiple studies on phenological changes in the pollen season have been published" (Lines 43-44).
Ziska et al. 2011 has been cited in the discussion of ragweed phenology in Section 4.2: "The apparent trend in the season end date for *Ambrosia* with PYAAT could be due to the increased number of frost-free days, consistent with global warming, and a strong relationship between frost-free days and changes of ragweed season length" (Lines 410-413).

Comment: Line 359: This line includes a reference to Table 2. There is no Table 2 in the manuscript. Should the reference be to Table 1?

Reponse: We apologize for this mistake. The table reference in section 4.3 should be to Table 1. This has been corrected.

Comment: Lines 505-507, section 5.2.4: This sentence, and this section, seems to neglect any mention that the first of the two observed ragweed peaks in the Mountain subregion (from about day 100 to day 140) is entirely missed by the model.

Response: We have added a note about the spring ragweed peak in the Mountain subregion in the discussion ("There is a yet unidentified observed spring peak of ragweed pollen at about day 125 in the Mountain subregion, possibly due to an identification error." lines 641-642), though no known source can yet be identified for these somewhat unusual ragweed pollen counts.

Comment: Note also that the lower row of plots in Figure 12 is mislabelled (except for the first in the row). What are currently r-u should really be q-t. q was missed somehow.

Response: Thank you for pointing out this error. Figure 12 has been updated with corrected panel labels, now including q and excluding u.

Comment: Line 532: See earlier comment regarding Parry et al. 2007. Also, a couple of additional references that would be strong support for this sentence are:
Lake IR, Jones NR, Agnew M, Goodess CM, Giorgi F, Hamaoui-Laguel L, Semenov MA, Solomon F, Storkey J, Vautard R, Epstein MM. Climate change and future pollen allergy in Europe. Environmental Health Perspectives 2017;125(3):385–391. DOI: 10.1289/EHP173
Ziello C, Sparks TH, Estrella N, Belmonte J, Bergmann KC, Bucher E, et al. Changes to airborne pollen counts across Europe. PLoS One 2012;7(4):e34076. DOI: 10.1371/journal.pone.0034076

Response: Additional references Ziello et al. 2012 and Lake et al. 2017 have been added to the citation in Line 664 as they support this discussion point. In addition, Lake et al. 2017 was cited in Lines 45-47 in the discussion of recent pollen dispersion modeling efforts.

Comments: Line 570: Change “estimating” to “estimation”.

Response: Change made on Line 703.

Comment: Line 584: Change “Association for” to “Academy of”.

Response: Change made on Line 718.

Comment: References: These should be carefully checked to ensure the details and format are correct. Details should be carefully checked against the PDF of each article.

Response: All references have been checked against their articles, and the details have been corrected for any references that were missing components or incorrectly formatted.

Comment: Line 731, Figure 1 caption: Instead of using the word “black” to describe the shading of the Pacific Northwest subregion, perhaps the term “dark grey” would be better.

Response: “black” updated to “dark grey” in Figure 1 caption as suggested.

Comment: Figure 2: The RAG and GRA lines are too similar. They are fine when enlarged on screen but when printed they are difficult to tell apart. Perhaps one could be red and the other black (meaning the four lines would be black, red, green, and blue).

Response: The line colors in Figure 2 have been updated such that the previously red ragweed line is now magenta. This appears to provide good contrast for all lines in this plot.

Comment: Line 733, Figure 2 caption: As stated earlier in the comments with respect to the manuscript as a whole, remove the word “climatological” and replace it with “average”, such as: Average daily observed time series of pollen count data . . .

Response: This language has been updated from “climatological” to “average”.

Comment: Line 740, Figure 3 caption: BELD is defined as “Biogenic Emissions Landuse Database” in Section 3.1 paragraph 2, not “Biogenic Emissions Land cover Database” as it is here in the figure caption. Which is correct?

Response: We apologize for this confusion. The correct name is “Biogenic Emissions Landuse Database”. The Figure 3 caption has been updated to reflect this.

Comment: Line 748, Figure 4 caption: Indicate that ragweed is “(g)”.

Response: Figure 4 caption has been updated accordingly.

Comment: Line 757, Figure 6 caption: Change the start of the caption to: Monthly average pollen emissions potential . . .

Response: This correction has been made to the Figure 6 caption.

Comment: Lines 777-782, Figure 12 caption: Change the three occurrences of climatologi-cal/climatology. The caption can start: Average daily (2003-2010) time series of pollen counts . . .

Response: The Figure 12 caption has been updated accordingly.

Comment: Line 778: Change RAG from “p-u” to “p-t”.

Response: The Figure 12 caption has been updated to reflect the changes to the panel labels in the figure.

Comment: Line 782: Add to the end of the very last sentence “by region and PFT”, i.e.: Note: scale of y-axes varies by region and PFT.

Response: This addition has been made to the Figure 12 caption.

Comment: Table 1: The numbers in the production factor (P) column should include the same number of numbers after the decimal point (I suggest 1, e.g., 89.1, 210.0, etc.), and the numbers should be aligned right in the column, not aligned left.

Response: All values right aligned. Decimal places are left unchanged for production factors as those are formatted to reflect the correct number of significant digits in that data. Linear regression values are also updated so that all numbers are rounded to two decimal places.

Best,

Matthew Wozniak

A prognostic pollen emissions model for climate models (PECM1.0)

Matthew C. Wozniak[1], Allison L. Steiner[1]

[1]Climate and Space Sciences and Engineering, University of Michigan, Ann Arbor, MI 48109, USA

Correspondence to: Matthew C. Wozniak (mcwoz@umich.edu)

Abstract. We develop a prognostic model of Pollen Emissions for Climate Models (PECM) for use within regional and global climate models to simulate pollen counts over the seasonal cycle based on geography, vegetation type and meteorological parameters. Using modern surface pollen count data, empirical relationships between prior-year annual average temperature and pollen season start dates and end dates are developed for deciduous broadleaf trees (*Acer, Alnus, Betula, Fraxinus, Morus, Platanus, Populus, Quercus, Ulmus*), evergreen needleleaf trees (Cupressaceae, Pinaceae), grasses (Poaceae; C_3, C_4), and ragweed (*Ambrosia*). This regression model explains as much as 57% of the variance in pollen phenological dates, and it is used to create a "climate-flexible" phenology that can be used to study the response of wind-driven pollen emissions to climate change. The emissions model is evaluated in a regional climate model (RegCM4) over the continental United States by prescribing an emission potential from PECM and transporting pollen as aerosol tracers. We evaluate two different pollen emissions scenarios in the model, using: (1) a taxa-specific land cover database, phenology and emission potential, and (2) a plant functional type (PFT) land cover, phenology and emission potential. The simulated surface pollen concentrations for both simulations are evaluated against observed surface pollen counts in five climatic subregions. Given prescribed pollen emissions, the RegCM4 simulates observed concentrations within an order of magnitude, although the performance of the simulations in any subregion is strongly related to the land cover representation and the number of observation sites used to create the empirical phenological relationship. The taxa-based model provides a better representation of the phenology of tree-based pollen counts than the PFT-based model, however we note that the PFT-based version provides a useful and "climate-flexible" emissions model for the general representation of the pollen phenology over the United States.

Matthew Wozniak 8/2/2017 5:14 PM
Deleted: PFT

Matthew Wozniak 8/2/2017 5:14 PM
Deleted: -

1 Introduction

Pollen grains are released from plants to transmit the male genetic material for reproduction. When lofted into the atmosphere, they represent a natural source of coarse atmospheric aerosols, ranging typically from 15 to 60 µm in diameter, while sometimes exceeding 100 µm (Cecchi 2014; Sofiev et al. 2014). In the mid-latitudes, much of the vegetation relies dominantly on anemophilous, or wind-driven, pollination (Lewis et al. 1983), representing a closely coupled relationship of pollen emissions to weather and climate. Anemophilous pollinators include woody plants such as trees and shrubs, as well as other non-woody vascular plants such as grasses and herbs. Pollen emissions are directly affected by meteorological (e.g., temperature, wind, relative humidity) and climatological (e.g., temperature, soil moisture) factors (Weber 2003). Aerobiology studies indicate that after release, pollen can be transported on the order of ten to a thousand kilometers (Sofiev et al. 2006; Schueler and Schlünzen 2006; Kuparinen et al. 2007) but there are still large uncertainties regarding emissions and transport of pollen.

Prognostic pollen emissions are useful for the scientific community and public, specifically for forecasting allergenic conditions or predicting the flow of genetic material. The interest and growing wealth of knowledge of allergenic pollen has been recently reviewed by *Beggs et al.* (2017). To date, most pollen emissions models focus on relatively short, seasonal time scales and smaller locales for a limited selection of taxa (Sofiev et al. 2013; Liu et al. 2016; R. Zhang et al. 2014). Climatic changes in large-scale pollen distributions are mostly absent from scientific literature, though multiple studies on phenological changes in the pollen season have been published (Ziska 2016; Yue et al. 2015; Y. Zhang et al. 2015a). Only recently have regional-scale modeling studies of pollen dispersion been conducted for Europe, and they have been used to assess the impacts of climate change on airborne pollen distributions (Sofiev and Prank 2016; Lake et al. 2017). In contrast to most meteorological pollen models, climate models require long-term (e.g., decadal to century scale) emissions at a range of resolutions covering continental regions up to the global scale. This distinction in both time and space requires a flexible model that can account for emissions without taxon-specific emission data (i.e. differentiation between genera or species) and can be used within aggregated vegetation descriptions, such as plant functional types (PFTs). Given recent interest in airborne biological particles and their role in climate (Despres et al. 2012; Myriokefalitakis et al. 2017), an emissions model that captures longer temporal scales and broader spatial scales is key to developing global inventories and understanding pollen's role in the climate system. Here we develop a model for use in the climate modeling community that can be used specifically to simulate pollen emissions on the decadal or centurial time scale for large regions using conventional climate or Earth system models.

Existing pollen forecasting models are often classified as either process-based phenological models or observation-based models (Scheifinger et al. 2013). Process-based phenological models employ a parameterization of plant physiology and climatic conditions (e.g., relating the timing of flowering to a chilling period, photoperiod, or water availability). Pollen season phenology in an anemophilous species is inherently connected to its environment via relationships in the growing season dynamics (e.g. bud burst and temperature, (Fu et al. 2012)), and many models apply the same techniques to flowering as for bud burst (Chuine et al. 1999). This approach to phenology could be suited to climate models, given its flexibility for adaptive phenological events and regional-scale studies. Typically, these types of phenological models are taxa specific as well as regionally dependent, e.g., *Betula* in Europe or

Matthew Wozniak 8/25/2017 12:15 PM
Deleted: from

Matthew Wozniak 8/25/2017 12:16 PM
Deleted: 10

Matthew Wozniak 8/9/2017 9:33 PM
Deleted: 70

Matthew Wozniak 8/25/2017 12:33 PM
Deleted: Mikhail

Matthew Wozniak 8/23/2017 11:24 AM
Deleted: M.

Matthew Wozniak 8/8/2017 12:27 PM
Deleted: Mikhail

ragweed in California (Sofiev et al. 2013; Siljamo et al. 2013; R. Zhang et al. 2014). These models are usually calibrated to local data only even though distinct geographic differences exist for pollen phenology. Thus, such models may not perform equally well in other locations. Though process-based models draw a connection between an atmospheric state variable, i.e. temperature, and pollen emissions, at least three parameters are required for optimization and they are susceptible to overfitting (Linkosalo et al. 2008). While some process-based models may be scaled up to larger regions while maintaining appreciable accuracy (García-Mozo et al. 2009), such models are generally not practical for implementation in larger-scale climate modeling with regional climate models (RCMs) and global climate models (GCMs) because sufficient land cover data is not available at the appropriate taxonomic level.

In contrast to process-based models, observation-based methods determine the phenology of vegetation with statistical-empirical approaches (e.g., relating the start of the pollen season with mean temperatures preceding the pollen event) and often rely on regression models or time series modeling (Scheifinger et al. 2013). Time series modeling utilizes observations to define the deterministic and stochastic variability of pollen count observations and is frequently used in aerobiological studies (Moseholm et al. 1987; Box et al. 1994). Regression models, either using a single or multiple explanatory variable(s), exploit past relationships to define the magnitude of emissions as well as timing variables such as the start date and duration of the pollen season (Emberlin et al. 1999; M. Smith and Emberlin 2005; Galán et al. 2008). Using local pollen count data, *Zhang et al.* (2015b) completed a regional phenological analysis using multiple linear regressions for pollen in Southern California for six taxa. *Olsson and Jönsson* (2014) show that empirical models based solely on spring temperature perform just as well as process-based models using the temperature forcing concept, and better than those including a chilling or dormancy-breaking requirement.

Observation-based methods assume stationarity, or the likelihood that the statistics of pollen counts or climate variables are not changing over time. For these models to apply outside of calibration period, they require that the driving pattern or relationship is maintained in the future (or past). For example, as the Earth's climate changes, these models do not represent the complex connections between pollen emissions and a warming world aside from the relationships determined empirically. However, these models provide clear and often simple formulations that have predictable behaviors and forgo the nuance of fitting ambiguous and uncertain parameters. We therefore choose to employ elements of the observational methods for this pollen emissions model formulation, as described in Section 4.

In addition to understanding the release of pollen grains, a second consideration is the large-scale transport of pollen. Once emitted to the atmosphere, pollen is mixed within the atmospheric boundary layer by turbulence, and depending on large-scale conditions, can be transported far from the emission source. Prior studies have used both Lagrangian (Hunt et al. 2002; Hidalgo et al. 2002) and Eulerian techniques to simulate the transport of pollen, with the former typically used for studies of crop germination and the latter primarily for allergen forecasting. For example, *Helbig et al.* (2004) used the meteorological model KAMM (Karlsruher Meteorologisches Modell) with the DRAIS (Dreidimensionales Ausbreitungs- und Immissions-Simulationsmodell) turbulence component to simulate daily pollen counts for region over Europe. *Schueler and Schlünzen* (2006) use a mesoscale atmospheric

model (METRAS) to quantify the release, transport and deposition of oak pollen for a two-day period over Europe. *Sofiev et al.* (2013) includes the long-range transport of birch pollen over Western Europe by developing a birch pollen map and a flowering model to trigger release in the Finnish System for Integrated modeling of Atmospheric coMposition (SILAM). *Efstathiou et al.* (2011) developed a pollen emissions model for use within the regional air quality model (the Community Multi-scale Air Quality model (CMAQ)), and tested their model with birch and ragweed taxa. *Zhang et al.* (2014) implements a similar pollen emissions scheme with a regional numerical weather prediction model (the Weather Research and Forecasting (WRF) modeling system). *Zink et al.* (2013) developed a generic pollen modeling parameterization for use with a numerical weather prediction model (COSMO-ART) that is flexible to include differing pollen taxa. Collectively, these relatively new developments suggest a growing interest in the prognostic estimation of pollen on the short-term for seasonal allergen forecasting on the weather (e.g., one to two weeks) time scale.

In this manuscript, we build on these coupled emissions-transport models and develop a comprehensive emissions model (Pollen Emissions for Climate Models; PECM) for use at climate model time scales that covers the majority of pollen sources in sub-tropical to temperate climes, including woody plants, grasses and ragweed. First, we summarize the spatial distribution and seasonality of pollen counts for various taxa in the United States based on current observations (Section 2). Then we develop new pollen emissions parameterization for climate studies (Section 4), transport these emissions over the continental United States (CONUS) using the Regional Climate Model version 4 (RegCM4) (Giorgi et al. 2012), and evaluate the results using eight years of observed pollen count data (Section 5). We implement two different land cover classification schemes to illustrate the uncertainties associated with vegetation representation for trees including: (1) detailed family- or genus- level tree distributions over CONUS, and (2) the use of plant functional type (PFT) level distributions, which groups vegetation types by physiological characteristics (Section 3). As the latter provides a greater opportunity for expansion into regional and global scale climate models over multiple domains, we discuss the effects that the PFT-based categorization has on the total estimated source strength of pollen. Finally, the limitations of this emissions framework and suggestions for future developments are included (Section 6).

2 Observed pollen Phenology

2.1 Data description

The National Allergy Bureau (NAB) of the American Academy of Allergy, Asthma and Immunology (AAAAI) conducts daily pollen counts at 96 sites in cities across the United States (US), its territories and several locations in southern Canada. All NAB sites implement a volumetric air sampler and certified pollen count experts to conduct daily pollen counts (grains m^{-3}) for up to 42 plant taxa at either the family level (e.g., Cupressaceae, Poaceae), genus level (e.g., *Acer*, *Quercus*), or for four generic categories termed "Other Grass Pollen," "Other Tree Pollen," "Other Weed Pollen" or "Unidentified." We use NAB pollen count data ranging from 2003-2010 at all stations in the continental United States (Figure 1) for selected taxa to develop and evaluate PECM, and to determine the phenology of wind-driven pollen. Individual station locations and descriptions are included in Table S1.

Matthew Wozniak 8/2/2017 5:43 PM
Deleted: Finnish emergency modeling system

Matthew Wozniak 8/2/2017 6:04 PM
Deleted: ; Table 1

We evaluate the observed pollen counts to determine the vegetation types that emit the largest magnitude of pollen over the continental United States. Since many of the taxa reported at the 96 NAB sites frequently have very low pollen counts (e.g., less than 10 grains m^{-3}), a threshold for the grain count is set to select the taxa with the highest pollen counts. We calculate the average of the annual maximum pollen count across all years (2003-2010), P_{avgmax}, at each site for each counted taxon. We then select taxa to include in PECM using two criteria: (1) the maximum of P_{avgmax} among all stations exceeds 100 grains m^{-3}, and (2) the average P_{avgmax} among all stations exceeds 70 grains m^{-3} (Table S2). Using these two criteria, 13 taxa are selected for inclusion in the model, including *Acer, Alnus, Ambrosia, Betula,* Cupressaceae, *Fraxinus,* Poaceae, *Morus,* Pinaceae, *Platanus, Populus, Quercus* and *Ulnus*. These thirteen taxa account for about 77% of the total pollen counted across the United States during 2003-2010.

The 13 dominant pollen types are grouped into four main categories by plant functional type: deciduous broadleaf forest (DBF), evergreen needle-leaf forest (ENF), grasses (GRA) and ragweed (RAG). Plant functional type is a land cover classification commonly used in the land surface component of climate models, and this categorization will allow flexibility to apply the emissions model to other climate models. The DBF category includes 9 genus-level taxa (*Acer, Alnus, Betula, Fraxinus, Morus, Platanus, Populus, Quercus,* and *Ulmus*) and the ENF category includes two family-level taxa (Cupressaceae and Pinaceae). The grass PFT utilizes pollen count data from the Poaceae family, although we note that the grass PFT classification may include herbs and other non-woody species that may emit pollen as well. *Ambrosia* (ragweed) is segregated as its own category (RAG), due to its high pollen counts in the early autumn and unique land cover features. Daily pollen counts were summed for each PFT prior to calculating an 8-year average pollen time series.

2.2 Observed seasonality of pollen emissions

Pollen counts are analyzed over five subregions based on their climatic differences (Figure 1; Table S1) to identify emissions patterns over the continental United States. These five subregions are the Northeast (temperate; 38°-48°N and 70°-100°W; 34 stations), the Southeast (temperate, subtropical; 25-38°N and 70°-100°W; 29 stations), Mountain (varied climate; 25°-48°N and 100°-116°W; 9 stations), California (Mediterranean, varied climate; 25°-40°N and 116°-125°W; 13 stations) and the Pacific Northwest (temperate rainforest; west of 116°W and north of 40°N; 4 stations). Figure 2 shows the observed average PFT daily pollen counts averaged over all stations within the defined subregions.

For deciduous broadleaf forest (DBF) taxa, the Southeast has the highest average pollen maximum reaching up to about 700-1200 grains m^{-3} around day 100. In the Northeast, DBF is the dominant PFT, reaching up to an average of 400 grains m^{-3} and peaking slightly later (around day 120) than the Southeast. California sites show an average peak around 150 grains m^{-3} occurring slightly earlier around day 80. A sharp maximum of 775 grains m^{-3} appears in the Mountain subregion at about day 80, with a secondary emission reaching around 150 grains m^{-3} on day 125. In the Northwest, DBF pollen has the earliest maximum (day 70) at about the same magnitude as California (~200 grains m^{-3}). In some locations, there is a secondary DBF peak in the late summer and early fall due to the late flowering of *Ulmus crassifolia* and *Ulmus parvifolia*, located predominantly in the Southeast and California (Lewis,

Matthew Wozniak 8/2/2017 6:07 PM
Deleted: e

Matthew Wozniak 8/2/2017 5:14 PM
Deleted: a climatological

Matthew Wozniak 8/9/2017 11:06 AM
Deleted: ; Figure 1

Matthew Wozniak 8/9/2017 10:55 AM
Deleted: include

Matthew Wozniak 8/9/2017 10:54 AM
Deleted: north of

Matthew Wozniak 8/9/2017 10:54 AM
Deleted: east of

Matthew Wozniak 8/9/2017 10:55 AM
Deleted: south of

Matthew Wozniak 8/9/2017 10:56 AM
Deleted: east of

Matthew Wozniak 8/9/2017 10:56 AM
Deleted: and

Matthew Wozniak 8/9/2017 10:57 AM
Deleted: west of

Matthew Wozniak 8/9/2017 10:57 AM
Deleted: and south of 40°N

Matthew Wozniak 8/2/2017 5:17 PM
Deleted: climatological

Matthew Wozniak 8/2/2017 5:17 PM
Deleted: climatological

Matthew Wozniak 8/2/2017 5:18 PM
Deleted: climatological maximum

et al.1983). In the Southeast this occurs between days 225 and 300, while in California this occurs twice around day 245 and day 265.

The two ENF families exhibit pollen release at two distinct but overlapping times, with Cupressaceae peaking before Pinaceae. Cupressaceae in the Southeast emits pollen earlier than in other subregions, with a maxima at just over 400 grains m^{-3} around day 10 and counts above 200 grains m^{-3} in December of the prior year. Cupressaceae dominates the total emissions for the Southeast, with a smaller maximum from Pinaceae of about 180 grains m^{-3} near day 110. In the Northeast, the bimodality of ENF is evident with the Cupressacaeae family reaching a maximum of 100 grains m^{-3} near day 85 with a secondary Pinaceae maximum approximately 65 days later at about half the magnitude (~50 grains m^{-3}). In the Mountain and Pacific Northwest subregions, the maximum occurs around day 50-80 and can reach up to 350 grains m^{-3} in the Mountain subregion, but in both subregions is generally much lower than the eastern United States (approximately 50 grains m^{-3}). In the California subregion, ENF emissions are comparatively low (< 50 grains m^{-3}) which is likely due to the bias in sampling locations.

The grasses (Poaceae) have comparatively low average pollen counts (<25 grains m^{-3}) throughout the season in all subregions except the Northwest, where the maximum reaches 75 grains m^{-3}. However, the average maximum Poaceae pollen count at individual stations is close to 100 grains m^{-3}, with the individual annual maxima reaching several hundreds of pollen grains m^{-3}. In the AAAAI data, there are two distinct maxima in the Northeast Poaceae count, and we attribute the first seasonal maximum to C_3 grasses (peak around day 155) and the second grass maximum mainly to C_4 grasses (peak around day 250). Observations by *Craine et al.* (2011) of Poaceae in an American prairie have indicated that C_3 and C_4 grass flowering occurs at distinctly different times, with C_3 in the late spring and C_4 in mid- to late summer. Similarly, *Medek et al.* (2016) observed two grass pollen peaks in Australia, with a stronger, late-summer peak at lower Southern latitudes where there is higher incidence of C_4 grass. However, the authors note that sometimes this may be due to a second flowering of some C_3 grass species. Although the C_3-C_4 separation cannot be confirmed in the AAAAI pollen count data because they are not distinguished during pollen identification, this distinction is included in the model as discussed in Sections 3.1 and 4.2 below. In the Southeast, this separation of the Poaceae pollen counts is less apparent because both of the emission maxima are broader and intersect one another. In the Southeast, the first observed pollen maximum (assessed as C_3 grass pollen) peaks earlier around day 140, while the second maximum (assessed as C_4 grasses) have a similar, yet smaller value around day 250. In the Mountain subregion, the first grass maximum occurs later in the year (day 175) and the second grass maximum occurs around day 250 in the late summer. Pollen counts in California are only substantial during the earlier flowering time (C3 grasses) and have a similar duration to the Northeast, peaking at around day 135. For the Pacific Northwest, there is one strong early peak of grass pollen in the middle of the summer (day 170) and a secondary maximum is negligible, although counts below 10 grains m^{-3} register around days 250-270.

Ragweed (*Ambrosia*) pollen is segregated from other grasses and herbs because of the strong allergic response in humans to this specific species and the unique timing of emissions. Because it is a short-day plant (i.e. its phenology driven by a shortening photoperiod and cold temperatures (Deen et al. 1998)), ragweed pollen seasons are generally constrained to the late summer with the exception of the Mountain region where some counts occur in the spring. Emissions in the Northeast reach a maximum around day 240 at 60 grains m^{-3} while they occur slightly

Matthew Wozniak 8/2/2017 5:18 PM
Deleted: climatological

Matthew Wozniak 8/7/2017 10:17 AM
Moved (insertion) [1]

Matthew Wozniak 8/7/2017 10:17 AM
Deleted: at

Matthew Wozniak 8/7/2017 10:17 AM
Deleted: at

Matthew Wozniak 8/7/2017 10:17 AM
Moved up [1]: In the AAAAI data, there are two distinct maxima in the Northeast Poaceae count, and we attribute the first seasonal maximum to C_3 grasses (peak at day 155) and the second grass maximum to C_4 grasses (peak at day 250).

Matthew Wozniak 8/7/2017 10:45 AM
Deleted:

Matthew Wozniak 8/7/2017 10:45 AM
Deleted: itself

later in the Southeast, peaking around day 270 with twice the magnitude (120 grains m^{-3}). Ragweed pollen in the Mountain subregion with an expected peak at around day 245, but also an earlier peak at around day 130 with no confirmed cause. *Ambrosia* is not detected in the station averages for California and the Pacific Northwest, although some individual sites in these regions record relatively low counts on the order of 10 grains m^{-3}.

Matthew Wozniak 8/2/2017 5:18 PM
Deleted: climatological

3 Model input data

3.1 Land cover data

With a goal of developing regional to global pollen emissions, one of the greatest limitations is the description of vegetation at the appropriate taxonomic level and spatial resolution. While land cover databases specific to species level are available for some regions, they are not available globally. Alternatively, vegetation land cover in regional to global models can be represented by classifications based on biophysical characteristics. For climate models, a common approach to represent land cover is with plant functional types (PFT), and global PFT data is readily available and used by many regional and global climate models to describe a variety of terrestrial emissions (Guenther et al. 2006) and biophysical processes in land-atmosphere exchange models. The creation of a pollen emissions model with PFT categorization would be of use at a broad range of spatial scales and domains while integrating more readily with climate models. In the pollen emissions model development and evaluation (Sections 4 and 5), we compare two different vegetation descriptions of broadleaf deciduous and evergreen needleleaf trees including (1) family- or genus-specific land cover and (2) land cover categorized by PFT.

The Biogenic Emissions Landuse Database version 3 (BELD) provides vegetation species distributions at 1 km resolution over the continental United States based on satellite imagery, aerial photography and ground surveys, as well as other land cover classification data such as geographical boundaries (Kinnee et al. 1997; https://www.epa.gov/air-emissions-modeling/biogenic-emissions-landuse-database-version-3-beld3). The BELD database includes 230 different tree, shrub and crop taxa across the United States as a fraction of the grid cell area at either the genus or species level. For family and genus level pollen emissions, the BELD land cover fraction for the 11 dominant pollen-emitting tree taxa identified in Section 2.1 is utilized (Table 1; Figure 3). For species level land cover data, land cover fraction is calculated as the aggregate of all species within a family or genus.

For the PFT land cover, we use the Community Land Model 4 (CLM4) (Oleson et al. 2010) surface dataset that employs a 0.05º resolution satellite-derived land cover fraction from the International Geosphere Biosphere Programme (IGBP) classification (Lawrence and Chase 2007). We sum all three biome PFT categories (temperate, tropical and boreal) for deciduous broadleaf forests (DBF) and two biome PFT categories (boreal and temperate) for evergreen needleaf forests (ENF) to produce the model PFT land cover.

Figures 4a-d compare the BELD land cover (summed by PFT) and CLM4 land cover for the two tree PFTs. Region by region comparison of land cover for all BELD taxa and each tree PFT (from both BELD and CLM4) is given in Table 2. An important distinction is that CLM4 land cover extends beyond U.S. borders because it is derived from a global dataset, whereas BELD is constrained to the continental United States. BELD and CLM4 land cover show

Matthew Wozniak 8/10/2017 3:03 PM
Deleted: CLM4

Matthew Wozniak 8/10/2017 3:06 PM
Deleted: IGBP

general agreement on the regional distribution of both tree PFTs. DBF is predominantly in the eastern portion of the United States with a gap in the Midwestern corn belt. ENF is present in the Southeast, the Northeast along the U.S.-Canadian border, along the Cascade and Coastal mountain ranges and throughout the northern Rockies. A notable difference is the CLM4 representation of ENF, which shows a strong, dense band extending from the Sierra Nevadas through the Canadian Rockies. The BELD ENF broadly covers the Rocky Mountain Range, yet more diffusely (land cover percentage up to 76%), whereas the CLM4 dataset shows sparser and dense ENF land cover (e.g., up to 100%) in the same range. For the DBF category, another notable difference is that the strong band of oaks around the Central Valley of California, which is evident in BELD but missing from the CLM4 data set. Additionally, the CLM4 has far greater densities of DBF along the Appalachian range than BELD. Overall, the CLM4 land cover fractions for forest PFTs are higher on average than the summed BELD taxa, about 2 to 10 times as much in each region, with the exception of California subregion DBF where CLM4 landcover is about half of that in the BELD dataset (Table 2).

Grass spatial distributions are given by C_3 (non-arctic) and C_4 grass PFT land cover classes from CLM4 (Figure 4e,f), which correspond to the observed family-level Poaceae pollen subdivided into C_3 and C_4 categories (described in Section 2.2). C_3 coverage is evident across the United States, with broad coverage throughout the Southeast, Midwest, and northern Great Plains (Fig. 4e). C_4 coverage is concentrated in the Southeast and Southern Great Plains at lower densities (Fig. 4f).

Ragweed requires a different land cover treatment, as land cover distributions are not available for ragweed across the entire continental United States. Ragweed is known to arise in areas of human disturbances (Forman and Alexander 1998; Larson 2003), and is found mainly in disturbed or developed areas such as cities and farms (Katz et al. 2014; Clay et al. 2006). *Ambrosia* land cover (Figure 4g) is derived from the urban and crop categories of the CLM4 land cover, which are sourced from LandScan 2004 (Jackson et al. 2010) and the CLM4 datasets, respectively. The urban data is subdivided by urban intensity, which is determined by population density. We assume that ragweed is unlikely to grow in the densest of urban areas (such as city centers), and utilize the lowest urban density category that is also the most widespread. Ragweed land cover (plants m^{-2}) in urban areas is determined by multiplying the average urban ragweed stemdensity given by *Katz et al.* (2014) by the urban land cover fraction. For crops, the CLM4 subdivides land cover fraction into categories including corn and soybean crops, and *Clay et al.* (2006) provide ragweed stem densities in soybean and corn cropland. Thus, we calculate the ragweed land cover in stems m^{-2} (f_{rag}):

$$f_{rag} = \alpha\left(d_{soy}f_{soy} + d_{corn}f_{corn}\right) + \beta\left(d_{urb}f_{urb}\right) \quad (1)$$

where d_{soy}, d_{corn} and d_{urb} represent the stem density (stems m^{-2}) of ragweed in soybean, corn and urban areas, respectively, and the f_{soy}, f_{corn} and f_{urb} represent the fractional land cover for soybean, corn and urban, respectively. α and β are tuning parameters to that are determined by a preliminary evaluation between modeled and observed ragweed pollen counts, where $\alpha = 0.01$ for crop and $\beta = 0.1$. *Zink et al.* (2017) show that a ragweed land cover representation developed by combining land use and local pollen count information evaluates better against observed pollen counts than even ragweed ecological models, giving confidence to this choice of land cover representation.

Matthew Wozniak 8/10/2017 3:06 PM
Deleted: IGBP

Matthew Wozniak 8/10/2017 3:06 PM
Deleted: IGBP

Matthew Wozniak 8/8/2017 12:47 PM
Deleted: more

Matthew Wozniak 8/10/2017 3:06 PM
Deleted: IGBP

Matthew Wozniak 8/10/2017 3:06 PM
Deleted: IGBP

Matthew Wozniak 8/10/2017 3:06 PM
Deleted: IGBP

Matthew Wozniak 8/10/2017 3:06 PM
Deleted: IGBP

Matthew Wozniak 8/10/2017 3:06 PM
Deleted: IGBP

Matthew Wozniak 8/9/2017 9:39 PM
Deleted: fraction

Matthew Wozniak 8/9/2017 9:39 PM
Deleted: LC_{rag}

All land cover data are regridded to a 25 km resolution across the United States to provide emissions at the same spatial resolution as the regional climate model (see Section 5).

3.2 Meteorological data for phenology

To develop the emissions model, we use two sources of meteorological data. The first is a high-resolution meteorological dataset to develop the phenological relationships for the timing of pollen release. Because reliable measurements are not available at all pollen count stations and there is uncertainty in the siting of these stations (e.g., they may be in urban areas with highly heterogeneous temperature), we use a gridded observational meteorological product for consistency across all sites (Maurer et al. 2002). The gridded Maurer dataset interpolates station data to a 1/8º grid across the continental United States on a daily basis, representing a high spatial resolution gridded data product where data from each meteorological station has been subject to consistent quality control. Higher resolution DayMet temperatures (daily 1 km) (Thornton et al. 2014) were used in lieu of Maurer data at NAB sites where the Maurer dataset did not provide information at the collocated grid cell (Table S1). For offline emission calculations input into the regional climate model, we use annual-average temperatures computed from monthly Climate Research Unit (CRU) temperature data (Harris et al. 2014). This data was interpolated from a 0.5ºx0.5º grid to the 25 km regional climate model grid used for pollen transport.

Matthew Wozniak 8/2/2017 7:12 PM
Deleted: met

4 PECM model description

4.1 Emission potential

The pollen emissions model is a prognostic description of the potential emissions flux of pollen (E_{pot}; grains m^{-2} d^{-1}) for an individual taxon *i*:

$$(2) \quad E_{pot,i}(x,y,t) = f_i(x,y)\frac{p_{annual,i}}{\int_0^{365}\gamma_{phen,i}(x,y,t)\,dt}\gamma_{phen,i}(x,y,t)$$

for a model grid cell of location *x* and *y* at time *t*. In this expression, *f(x,y)* is the vegetation land cover fraction (Section 2.1; m^2 vegetated m^{-2} total area), p_{annual} is the daily production factor (grains m^{-2} yr^{-1}), and γ_{phen} is the phenological evolution of pollen emissions that controls the release of pollen (description below). Equation 2 can apply to either a single taxa or PFT, depending on the prescription of land cover through *f(x,y)*. In the simulations described here, emissions are calculated offline based on this equation and provided as input to a regional climate model (RCM). This emission potential is later adjusted based on meteorological factors in the RCM where the pollen grains are transported as aerosol tracers (Section 5.1.1). In the future, Equation 2 can be coupled directly within the climate model for online calculation of emissions. The phenological and production factors are described in greater detail below.

4.2 Phenological factor (γ_{phen})

Based on the observed pollen counts, a Gaussian distribution is used to model the phenological timing of pollen release (γ_{phen}):

$$(3)\quad \gamma_{phen,i}(x,y,t) = e^{-\frac{(t-\mu(x,y))^2}{2\sigma(x,y)^2}}$$

where $\mu(x,y)$ and $\sigma(x,y)$ are the mean and half-width of the Gaussian, respectively, and can be determined based on the start day-of-year (sDOY) and end day-of-year (eDOY) calculated by an empirical phenological model:

$$(4)\quad \mu(x,y) = \frac{sDOY(x,y) + eDOY(x,y)}{2}$$

$$(5)\quad \sigma(x,y) = \frac{eDOY(x,y) - sDOY(x,y)}{a}$$

The fit parameter, a, accounts for the conversion between the empirical phenological dates based on a pollen count threshold and the equivalent width of the emissions curve. Based on evaluation versus observations, a = 3 was selected for initial offline simulations.

Linear regressions of observed sDOY and eDOY from individual pollen count stations versus temperature are used to empirically determine sDOY and eDOY that drive γ_{phen}. An important criteria is the grain count used to determine the sDOY and eDOY, and we utilize a count threshold adaptable to bimodal emission patterns such as those noted for *Ulmus* and Poaceae. *Sofiev et al.* (2013) selected dates on which the 5th and 95th percentile of the annual index (annual sum of pollen counts) were reached, while *Liu et al.* (2016) combined a 5 grains m^{-3} threshold with the additional condition that 2.5% (97.5%) of the annual sum of pollen was reached before the start (end) date. Here, we implement a pollen count threshold of 5 grains m^{-3} and found this was sufficient to reproduce the observed seasonal cycle. To account for smaller signals that may be due to count errors (e.g., an exceedance of the 5 grains m^{-3} threshold but not followed by an increase in emissions), we used a moving window with a threshold of 25 grains m^{-3} for the sum of pollen counts in the nearest 10 neighboring days; when the sum of the neighbors failed to meet this threshold, the data point was omitted. In this manner, we calculated the sDOY and eDOY for the full 8-year time series for each taxon at each station. If more than one start or end date was found in a single year at a single station for a taxon that was not clearly bimodal, only the first set of dates was retained for the linear regression. For taxa with an observed bi-modal peak, the second peak was treated as a separate taxon (e.g. early and late *Ulmus*, C_3 and C_4 Poaceae) with a separate phenology. Once the sDOY and eDOY were determined, outliers in these dates were determined by bounding the data for each taxon at four times the mean absolute deviation of sDOY and eDOY.

Near surface atmospheric temperature (e.g., 2m height) is an important factor of vegetation phenology. In the interest of having a regional model of emissions that prognostically calculates the start dates, the previous year annual average temperature (PYAAT) based on near-surface atmospheric temperature from *Maurer et al.* (2002) and *Thornton et al.* (2014) (Section 2.2) is the explanatory variable in the linear regressions. For example, for a start date of February 2, 2007, the PYAAT would be the mean temperature for the year 2006. For *Pinus* and Cupressaceae, PYAAT is calculated differently from July 1, 2005 - June 30, 2006 because emissions of these families begin in the early winter (December). Prior studies have shown that the meteorology of the year previous to the pollen season influences pollen production, especially temperature, suggesting that PYAAT may be a good

Matthew Wozniak 8/25/2017 1:31 PM
Deleted: -

Matthew Wozniak 8/25/2017 1:31 PM
Deleted: -

predictor variable (Menzel and Jochner 2016). While emissions in this study are calculated using offline meteorological data, this also could be coupled to a dynamic land surface model to predict reasonably accurate pollen phenological dates.

To exemplify this method, Figure 5 shows the phenological dates and regression lines for the *Betula* (birch) genus, with all 13 modeled taxon shown in Figures S1 and S2. The sDOY and eDOY of the pollen season show a moderate and considerable trend with temperature for most taxa and PFTs (Table 1; Figures S1 & S2). The linear regression models for sDOY explain 41% of the variance on average for DBF taxa, 47% on average for ENF taxa, 48% for C_3 Poaceae, and 8% for *Ambrosia* while having a negligible R^2 for C_4 Poaceae. For eDOY, the linear regression models explain 21% of the variance on average for DBF taxa, 29% for ENF, 4% for C_3 Poaceae, 32% for C_4 Poaceae, and 37% for *Ambrosia*. All trends except C_4 Poaceae, late elm, and *Ambrosia* are negative, indicating that warmer previous-year temperatures result in earlier start and end dates. For most tree taxa, the trend of both sDOY and eDOY are negatively correlated with PYAAT, with a steeper negative slope for sDOY. The correlation for the duration of the pollen season (eDOY – sDOY) is then positive for all taxa except Cupressaceae. This suggests that warmer climates have earlier pollen season start and end dates but longer season lengths.

Trends for grass in Australasia show that the correlation of the end date of the pollen season with average spring temperature is positive, while the same relationship for the start date is negative, suggesting also that season start dates are earlier and season duration increases with warmer climates (Medek et al. 2016). The apparent trend in the season end date for *Ambrosia* with PYAAT could be due to the increased number of frost-free days, consistent with global warming and a strong relationship between frost-free days and changes of ragweed season length (Easterling 2002; Ziska et al. 2011).

This agrees with earlier findings that suggest the pollen season will, on average, start earlier with a warmer global climate and have a longer duration (Confalonieri et al. 2007). The spatiotemporal heterogeneity of climate change may affect which regions and seasons will be most influenced by climate change (Ziska 2016). In fact, there is imperfect agreement that earlier start dates and longer seasons will occur unanimously throughout the United States region, at least for trees (Yue et al. 2015). It is understood that photoperiod and the dormancy-breaking process controlled by chilling temperatures play a significant role in the phenology of trees (Myking and Heide 1995; Ziska 2016), and it is generally accepted that a plethora of other factors, such as plant age, mortality, and nutrient availability also affect observed phenological dates (Jochner et al. 2013). However, even without these factors, the current phenological model is applicable to large regions and provides a clear response of plants to inter-annual climate variability as well as long-term climate changes. For this first assessment of PECM, we assume that the pollen production factor (p_{annual}) does not change with time and that the phenological model described above captures the main features of pollen emissions.

4.3 Annual pollen production (p_{annual})

Annual production factors (grains m^{-2} $year^{-1}$, where m^{-2} refers to vegetated area, or grains $stem^{-1}$ $year^{-1}$ for ragweed) for each modeled taxon are provided in Table 1. The annual pollen production factor (p_{annual}) defines the amount of pollen produced per vegetation biomass per year based on literature values. *Tormo Molina et al.* (1996) report the

Matthew Wozniak 8/25/2017 1:30 PM
Deleted:

Matthew Wozniak 8/16/2017 2:33 PM
Deleted: A shows

Matthew Wozniak 8/16/2017 2:36 PM
Deleted: ,

Matthew Wozniak 8/16/2017 3:13 PM
Deleted: the

Matthew Wozniak 8/16/2017 2:38 PM
Deleted: of other empirical modeling studies

Matthew Wozniak 8/7/2017 1:52 PM
Deleted: Though such an outcome is more intuitive,

annual pollen productivity in grains tree^{-1} year^{-1} measured from three representative trees from several taxa. *Morus* has no known reference for production factor and was assumed to be 10x10^{7} grains m^{-2} year^{-1}, conservatively at the low end of the range for other deciduous broadleaf taxa. Other tree taxa and grasses are reported in grains m^{-2} year^{-1}, while ragweed is reported in grains stem^{-1} year^{-1} (Helbig et al. 2004; Jato, Rodríguez-Rajo, and Aira 2007; Hidalgo, Galán, and Domínguez 1999; Prieto-Baena et al. 2003; Fumanal, Chauvel, and Bretagnolle 2007). To convert the production factors from *Tormo Molina et al.* (1996) (grains tree^{-1} year^{-1}), the production factors for each representative tree are multiplied by the tree crown area, calculated as the circular area of the tree crown diameter given in Table II of *Tormo Molina et al.* (1996). The resulting individual production factors (grains m^{-2} year^{-1}) are then averaged for each taxa.

After sensitivity experiments of running pollen emissions in RegCM4, we find that the literature value of p_{annual} for Poaceae provides better agreement with observations for C_4 grass when reduced by a factor of 10, thus we use this value. To obtain the coefficient of daily pollen production over the duration of the phenological curve, γ_{phen}, the integral of the daily pollen production is normalized to p_{annual} as demonstrated by Equation 2.

4.4 Offline emissions simulations

We calculate emissions offline for two versions of PECM that differ in the land cover input data for woody plants. The first uses the detailed BELD tree database (Figure 3) for tree pollen emissions (hereinafter the "BELD" simulation), and the second uses globally based PFT data for tree pollen emissions (Figures 4b and 4d) (hereinafter the "PFT" simulation). For the grass and ragweed taxa, the emissions calculations are identical between the two simulations as the input land cover is the same for these two categories. While the family and genus level is useful for the allergen community, the respective taxon land cover databases needed to develop a global, adaptable model are not always available. While many plant traits are found to vary quite strongly within individual PFTs (Reichstein et al. 2014), the PFT convention is accepted and remains in use in climate models, particularly because of the lack of species-level land cover data at large scales. For the PFT version, pollen counts from individual taxa were summed within each PFT prior to calculation of the phenological regression (Table 1). We exclude the bimodality in *Ulmus* for the PFT version because it is the only tree taxon that exhibits this behavior, and late *Ulmus* pollen emissions are relatively small compared to the major DBF season. The production factors for each PFT are calculated as the unweighted average of the production factors for all the taxa within the PFT (Table 1).

Figures 6-11 show the monthly averages of the 2003-2010 emissions potential calculated by the offline models described in Section 4.1 (E_{pot}; Equation 2). The seasonal cycle can be clearly identified in the emissions potential, with the onset of pollen emissions beginning in the warmer south and moving northward along the gradient of annual average temperature. Colder locales such as those at high elevations can interrupt this general trend. Though pollen seasons generally end later in the colder parts of the domain just as they start later, modeled pollen emission seasons tend to be shorter at colder locations for most taxa (about 1 day per 1ºC, on average). The highest maximum emissions for DBF occur over the Appalachian range between April and May for both the BELD and PFT versions (Figures 6 and 7). For ENF, the maximum occurs in April in the American West for the BELD version where Cupressaceae land cover is dominant, while it is consistent in magnitude between the Southeast and West Coast for

Matthew Wozniak 8/25/2017 12:59 PM
Deleted: a number of

Matthew Wozniak 8/25/2017 1:00 PM
Deleted: .

Matthew Wozniak 8/10/2017 1:53 PM
Deleted: The

Matthew Wozniak 8/23/2017 11:39 AM
Deleted: similar to

Matthew Wozniak 8/17/2017 8:57 PM
Deleted: other tree taxa, grasses and ragweed are reported in either grains

Matthew Wozniak 8/9/2017 9:42 PM
Deleted: per tree or stem, or grains per unit vegetated area

Matthew Wozniak 8/17/2017 8:55 PM
Deleted: All values are converted to an annual production factor (grains m^{-2} year^{-1}) for each modeled taxon (Table 2).

Matthew Wozniak 8/25/2017 1:03 PM
Deleted: Tormo Molina et al.

Matthew Wozniak 8/17/2017 9:00 PM
Deleted: for C_4 grass

Matthew Wozniak 8/17/2017 9:00 PM
Deleted: in

Matthew Wozniak 8/25/2017 1:04 PM
Deleted: climatology

the PFT-based version (Figures 8 and 9). The grass PFT maximum emissions occur in June in the northern Rockies for C_3 and in September in the South-Central Great Plains for C_4 (Figure 10). Ragweed pollen emissions reach their maximum during September throughout the Corn Belt where soybean and corn crops dominate the land surface, with local maxima apparent in urban centers (Figure 11).

5 Emissions implementation and evaluation

5.1 Emissions implementation in a regional climate model

To evaluate PECM, emissions calculated offline are included within a regional climate model to compare simulated atmospheric pollen concentrations with ground-based observations from the NAB pollen network. The two phenological pollen emissions estimates (BELD and PFT) described above are prescribed as daily emissions, after which they are scaled by meteorological factors and undergo atmospheric transport. We use the Regional Climate Model version 4 (Giorgi et al. 2012), which is a limited-area climate model that includes a coupled aerosol tracer module (Solmon et al. 2006) that readily accommodates pollen tracers (Liu et al. 2016). The pollen tracer transport scheme is extended from one to four bins in this study to simulate the four PFTs (DBF, ENF, GRA, and RAG), with tracer bin particle effective diameters of 28 µm, 40 µm, 35 µm and 20µm, respectively. Additionally, the temporal emissions input is updated to accommodate daily pollen emissions (grains m^{-2} day^{-1}).

RegCM4 is based on the hydrostatic version of the Penn State/NCAR mesoscale model MM5 (Grell et al. 1994) and configured for long-term climate simulations. In our RegCM4 configuration, we use the Community Land Model version 4.5 (CLM4.5; (Oleson et al. 2010)), the Emanuel cumulus precipitation scheme over land and ocean (Emanuel 1991), and the SUBEX resolvable scale precipitation (Pal et al. 2000). The horizontal resolution is 25-km with 144x243 grid cells on a Lambert Conformal Projection centered on 39ºN, 100ºW with parallels at 30ºN and 60ºN (Figure 1). The vertical resolution includes 18 vertical sigma levels. Boundary conditions are driven by ERA-Interim Reanalysis while sea surface temperatures are prescribed from NOAA Optimum Interpolation SSTs (Dee et al. 2011; Smith et al. 2008). Two 8-year simulations of pollen emissions and transport in RegCM4 were conducted from 2003-2010 with the BELD and PFT version of the offline emissions model. Six months of spin-up (July-December 2002) are run for both simulations that we exclude from the following analysis.

In the model, we calculate the fate of four pollen tracers corresponding to the four PFTs (DBF, ENF, GRA and RAG) from the PECM offline emissions. Because individual tracers add to the computational cost of the simulations, BELD-based tree emissions are summed into DBF and ENF PFTs before they are emitted into the model atmosphere. To calculate the emissions, the emission potential calculated offline for each PFT (E_{pot}) is scaled according to surface meteorology following the methods of *Sofiev et al.* (2013):

(5) $E_{pollen,i}(x,y,t) = E_{pot,i}(x,y,t) f_w f_r f_h$

(6) $f_w = 1.5 - e^{-(u_{10}+u_{conv})/5}$

Matthew Wozniak 8/9/2017 9:46 PM
Deleted: and

$$(7)\quad f_r = \begin{cases} 1, & pr < pr_{low} \\ \dfrac{pr_{high} - pr}{pr_{high} - pr_{low}}, & pr_{low} < pr < pr_{high} \\ 0, & pr > pr_{high} \end{cases}$$

$$(8)\quad f_h = \begin{cases} 1, & rh < rh_{low} \\ \dfrac{rh_{high} - rh}{rh_{high} - rh_{low}}, & rh_{low} < rh < rh_{high} \\ 0, & rh > rh_{high} \end{cases}$$

where f_w, f_r, *and* f_h are the wind, precipitation and humidity factors, respectively. The meteorological parameters in these equations are from online RegCM variables, including u_{10} and u_{conv} as the 10-meter horizontal wind speed and vertical wind speed, and *pr* and *rh* are precipitation and relative humidity with low and high thresholds. These scaling factors account for the effects of wind, precipitation and humidity on the emission of pollen from flowers and cones. The humidity and precipitation factors are piecewise linear functions of the near-surface (10 m) RH and total precipitation and range from 0 (high precipitation or humidity) to 1 (no precipitation or low humidity). The wind factor ranges from 0.5 to 1.5, as even in calm conditions turbulent motions can trigger pollen release with high winds releasing more pollen. These scaled emissions are then transported according to the tracer transport equation (Equation 9) of *Solmon et al.* (2006) that includes advection, horizontal and vertical diffusion (F_H and F_V), convective transport (T_c), as well as wet (R_{Wls} and R_{Wc}, representing large scale and convective precipitation removal) and dry deposition (D_d) of an individual tracer (χ), represented by i = 1 to 4 for each PFT pollen emission:

$$(9)\quad \frac{\partial \chi^i}{\partial t} = \bar{V} \cdot \nabla \chi^i + F^i{}_H + F^i{}_V + T^i{}_C + S^i - R^i{}_{Wls} - R^i{}_{Wc} - D^i{}_d$$

5.1 Model evaluation against observations

We evaluate the efficacy of PECM in simulating the timing and magnitude of pollen emissions across the continental United States by evaluating RegCM4 tracer concentrations versus observations. We compare the average daily simulated near-surface pollen counts and observed, ground-based pollen counts for each of the four modeled PFTs (Figure 12). The observed pollen time series in Figure 12 are the spatial average of the average daily pollen counts at all pollen counting stations comprising each of the five major U.S. subregions (Section 2.2) and are compared with the modeled average daily pollen counts, which averages the individual grid cells that contain the pollen counting stations. Interannual variability is assessed using the relative mean absolute deviation for each day of the average time series. The inter-annual variability in observed daily pollen counts throughout the year is, on average, 81, 78, 78 and 77% of the mean (DBF, ENF, grass and ragweed, respectively), while this variability from the simulations is 53% for the BELD version of the DBF model and 61% for the PFT version, 55% and 92% for the BELD and PFT versions of the ENF model, 43% for grasses, and 49% for ragweed (Figure 12). This indicates that the model is capturing the relative inter-annual variability of the pollen counts between PFTs, but not all of the variability in pollen counts from season to season. The unexplained variability in pollen concentrations could be due to the lack of sensitivity of annual pollen production factor to the environment, as this may be closely tied with precipitation (Duhl et al. 2013) or temperature (Jochner et al. 2013). Additionally the average observed and simulated pollen counts are analyzed using box-and-whisker plots to assess the models' representivity of pollen

Matthew Wozniak 8/2/2017 5:20 PM
Deleted: daily climatology of

Matthew Wozniak 8/2/2017 5:24 PM
Deleted:

Matthew Wozniak 8/2/2017 5:21 PM
Deleted: climatological

Matthew Wozniak 8/2/2017 5:21 PM
Deleted: daily climatology

Matthew Wozniak 8/25/2017 1:07 PM
Deleted: climatology

Matthew Wozniak 8/2/2017 5:25 PM
Deleted: climatology

Matthew Wozniak 8/2/2017 5:26 PM
Deleted: climatological

Matthew Wozniak 8/2/2017 5:26 PM
Deleted: is

count magnitude in spite of phenology (Figure 13). These metrics are discussed in detail by PFT and U.S. subregion below.

5.2.1 DBF

In the Northeast, the BELD model captures both the observed seasonal timing and the magnitude of DBF pollen counts (Figure 12a). Observed DBF phenology is also simulated by the PFT-based emissions with even greater statistical accuracy in reproducing the observed pollen counts, though the BELD model more accurately reproduces the annual maximum (Figure 13a). The accuracy in this subregion is not surprising, as Northeastern pollen counting stations contributed the greatest number of data points to the phenological regression analyses. Observed DBF pollen counts in the Southeast have a large maximum that is greater than the average seasonal maximum of all four other subregions and all three other PFTs (Figure 12b), which is predominantly from *Quercus*. Neither the BELD nor PFT version of the simulation recreates this sharp peak, but they do simulate a large majority of the pollen count distribution (Figure 13b), especially the PFT-based model for which the lower 75% of simulated average pollen counts agrees well with the lower 75% of observed average pollen counts. The PFT model does not specifically resolve *Quercus*, and while the BELD model does resolve *Quercus*, it fails to model this maximum. This may be because the linear regression producing the phenological dates is an average, where a longer season may result from earlier start dates and/or later end dates that will reduce the maximum of the Gaussian distribution of pollen counts in the time series. In the Mountain region, there is an observed maximum early in the spring that is not simulated by either model because the DBF phenology at several cold Mountain sites is exceptionally early, and falls well below the regression lines (Figures S1, S2). However, both the BELD and PFT model simulate the second Mountain subregion peak with the correct magnitude. The BELD simulated maximum DBF in California is about 40 days later than the observed peak, also due to the regionally anomalous phenology in California as compared with the rest of the U.S., and though the PFT model peaks much closer to the observations, it underestimates DBF pollen counts. In the Pacific Northwest, the observed pattern is quite similar to the DBF pollen phenology in the Mountain subregion with only a slightly weaker early spring peak due to low-elevation pollen. The observed phenological pattern (Fig. 12e) and pollen count magnitudes (Fig. 12e) are both more accurately simulated by the BELD model, likely due to the earlier spring maximum that does not appear in the PFT simulation.

5.2.2 ENF

Like DBF, the BELD ENF in the Northeast is well represented by simulating two distinct Cupressaceae and Pinaceae maxima, although the model slightly underestimates observed Pinaceae pollen counts (Figure 12f). The PFT model ENF phenology emits from the start of the earlier Cupressaceae season to the end of the later Pinaceae season, while overestimating the maximum pollen count by about a factor of 2. In the Southeast, the winter peak is not captured by the model phenology (Figure 12g). However, the spring Pinaceae maximum is accurately captured by the BELD simulation. The PFT model follows the observed Pinaceae phenology more closely, though overestimating pollen counts by a factor of 2 to 3 and estimating a later ending date by about 40 days. In the Mountain subregion, ENF start and end dates are simulated by the BELD model with improved accuracy than the

Matthew Wozniak 8/2/2017 5:26 PM
Deleted: climatological

Matthew Wozniak 8/2/2017 5:27 PM
Deleted: climatological

Matthew Wozniak 8/2/2017 5:27 PM
Deleted: climatological

Matthew Wozniak 8/8/2017 1:01 PM
Deleted: which

Matthew Wozniak 8/10/2017 2:28 PM
Deleted: maxima in

Matthew Wozniak 8/10/2017 2:48 PM
Deleted: due to negligible BELD land cover fractions for Cupressaceae (Figure 4c)

DBF phenology in this subregion, though the predicted spring maximum is later than observed (Figure 12h). As with DBF, there is good agreement between the BELD model with the later part of the season in this subregion. The PFT model, again, simulates the peak ENF emissions in the later part of the season and overpredicts the pollen counts by a factor of 2 to 3. In the California subregion, the tails of the pollen distributions by both models closely resemble the pollen count magnitudes, yet the majority of these pollen counts (the top 75%, Figure 13i) lie above the observed maximum (Figure 12i). Finally, in the Pacific Northwest, the BELD model phenology shows some agreement with the model mean (Figure 13j), with the simulated pollen count showing a stronger Gaussian distribution than observed (Figure 12j). In contrast, the PFT model grossly overpredicts the observed pollen counts by up to a factor of 10 at its maximum, likely due to the greater representation of the ENF PFT than the BELD model in this region. The simulated average start date of the PFT model is within a few days of the observed average start date, while the end date is about 20 days later than observed.

Matthew Wozniak 8/2/2017 5:27 PM
Deleted: climatological

Matthew Wozniak 8/2/2017 5:27 PM
Deleted: climatological

5.2.3 Grasses

Grass phenology across all subregions for both C_3 and C_4 types is captured by the emissions estimates (Figure 12 k-o). However, the pollen count magnitude in Northeastern C_3 grass peak is overestimated by about a factor of seven, even when using the minimum value of the annual production factor in the range estimated by *Prieto-Baena* (2003) (Figure 12k). The secondary peak, which we attribute to C_4 grasses and is only about half as large, is well-represented. In the Southeast, the simulated pollen count magnitudes are much closer to observations, while the C_3 peak is overestimated here by only a factor of 2 and the C_4 peak is within 5 grains m^{-3} (Figure 12l). In this region, the observed duration of the pollen emissions is not fully captured by the simulated grass phenology in the Southeast, and this is probably due to the non-Gaussian shape of the observed time series. In the Mountain subregion, the C_3 pollen count is overestimated by the model, but the phenology is represented by a gradual rise in low emissions beginning in March to match the maximum burst of emissions in June (Figure 12m). C_4 grass pollen counts are not simulated in the Mountain region due to the relatively low C_4 land cover in the CLM4 dataset (Figure 4f). In California there is a single observed grass peak, which the model attributes to C_3 pollen, and the peak count in the simulation is about 5 days late and about 2 to 10 times too large (Figure 12n). In the Pacific Northwest, the average C_3 season is accurately simulated with the exception that the phenology is shifted 20 days earlier than observed (Figure 12o). A small C_4 peak in the observations at around day 260 is not simulated in this region due to negligible land cover for C_4 grasses in the CLM4 land cover data (Figure 4f).

Matthew Wozniak 8/10/2017 3:06 PM
Deleted: IGBP

Matthew Wozniak 8/2/2017 5:27 PM
Deleted: climatological

Matthew Wozniak 8/10/2017 3:06 PM
Deleted: IGBP

5.2.4 Ragweed

Simulated ragweed phenology in the Northeast, Southeast, and Mountain subregions follows the observed phenology of late-summer ragweed very closely, where the peaks of both the simulated and the observed time series averages occur within a day of each other (Figures 12p-t). Close evaluation of each regional phenological time series reveals that many of the observed features, like those determined by the rate of increase or decrease of the pollen count, are reproduced by the model. The magnitude of the modeled ragweed maxima in the Northeast and Mountain subregions is slightly greater than observed (Figures 12p and 12r), while there is a clear underestimation by a factor

Matthew Wozniak 8/2/2017 5:28 PM
Deleted: climatology

of 4 or 5 in the Southeast (Figure 12q). There is a yet unidentified observed spring peak of ragweed pollen at about day 125 in the Mountain subregion, possibly due to an identification error. The observed average ragweed pollen counts in California and the Pacific Northwest are negligible, though the simulation predicts them to be similar in magnitude and timing to the other three subregions (Figure 12s and 12t). These discrepencies may be due to the land use description developed for ragweed (Section 3.1), which may overestimate the ragweed potential in the western United States, or potentially the relatively spare observational stations in these regions may be poorly placed relative to emissions sources.

Matthew Wozniak 8/2/2017 5:28 PM
Deleted: climatological

6 Conclusions

We have developed a climate-flexible pollen emissions model (PECM) for the 13 most prevalent wind-pollinating taxa in the United States based on observed pollen counts. PECM was adapted to the PFT categorization common to climate and Earth system models with four major temperate-zone PFTs (DBF, ENF, grasses and ragweed), thus it is possible to apply this model to larger geographic regions where specific taxon-level data is unavailable. We evaluated PECM using a regional climate model (RegCM4) to transport emissions and evaluated resulting pollen counts versus observations. PECM generally captures the observed phenology, and observed surface pollen concentrations can be simulated within an order of magnitude. While many emissions models to date have focused on smaller geographical regions with more detailed land cover information and pollen information, this model represents the first of its kind to simulate multiple taxa over broad spatial areas. This transition to a larger scale does have its disadvantages, and we define several major sources of uncertainty to consider when scaling up pollen emissions to the regional or global scale: (1) pollen production factors, (2) climatic sensitivities in phenological timing, (3) land cover data, and (4) taxa specificity. We discuss each of these uncertainties in greater detail.

A large source of uncertainty is the use of a constant annual production factor for pollen (Section 4.3). It has been reported that wind-driven pollen production has increased historically and is expected, potentially, to increase in the near future (R. Zhang et al. 2014; Lake et al. 2017; Confalonieri et al. 2007; Ziello et al. 2012). Some of more effective improvements to the emission model would be to create a pollen production model that is sensitive to multiple environmental factors such as soil moisture, temperature and nutrient status (Jochner et al. 2013). The interannual variability in observed daily pollen counts is, on average, substantially greater than that of the modeled pollen counts, which is likely due to this lack of production sensitivity. The current production factors for woody plants could be enhanced by studies that extend the number of representative units (i.e. individual trees) of vegetation used to determine the average pollen production. In a PFT representation, there is an inevitable limitation to the accuracy of any single PFT's ability to account for taxa differences within the PFT. Furthermore, the current model also assumes that there are no interspecies differences that affect the performance of the BELD model as well as the PFT model, whereas in reality it may vary by an order of magnitude within a genus (Duhl et al. 2013). However, despite the assumption of a constant production factor, observed surface pollen counts for all PFTs are typically reproduced within a single order of magnitude, as apparent in emission model evaluation.

Second, the use of observed relationships between pollen count and temperature to determine the phenological pollen start and end date also adds uncertainty to our modeling framework. Firstly, we assume stationarity in the phenological relationships, and this assumption may be violated. Secondly, based on the subregions defined for the analysis, there appears to be a bias in the linear regressions toward subregions with more available pollen counting stations, therefore affecting performance differences in these regions. Lastly, even though generally the Gaussian time series model of the pollen phenology performs well in our analysis, in the PFT representation the Gaussian absorbs or misses some of the phenological details in the observed pollen seasonality, and in some cases taxa (e.g. grasses in the Southeast subregion) may not be captured by the existing phenology.

Third, the specificity of land cover data provides an important constraint in the overall simulation of emissions. The representation of land cover is a key factor to accurately capturing regional features, especially in areas with a high degree of topographical variation and therefore greater variance in the land cover. For example, we notice large differences in the two model simulations when considering tree-specific taxa, such as in the western United States for ENF (Section 5.2.2). Also, our definition of the land cover available for ragweed used assumptions based on crop cover and urban area, which overestimated emissions in the western United States (Section 5.2.4). Interestingly, even though ragweed lacks an exact spatial distribution, distinct observed features of the ragweed phenology in three of the five subregions emerged using the current ragweed land cover parameterization.

Fourth, the aggregation of emissions to the PFT level affects the representativeness of the production factors, phenology and land cover. When comparing the two models of the tree pollen (BELD versus PFT), the individual phenology of each of the 11 tree taxa are resolved by the BELD simulation, whereas they are either folded into or excluded from the single phenology modeled by the PFT simulation. This results from either treating the taxa in the phenological regressions individually, as in the BELD model, or as a sum, as in the PFT model. With a few exceptions (e.g., the ENF family distinctions), the PFT model does generally reproduce the regional phenology throughout the United States domain, which is a priority of this study.

Despite these limitations, the empirical formulation presented here is the first of its kind to predict a broad range of different pollen emissions across a large geographic region. Even with univariate phenology and invariable pollen production factors, the model includes seasonal dynamics sensitive to climate change consistent with observations and is also able to simulate observed pollen magnitudes. As a result, the model can be useful for estimation of how allergenic risk or plant reproductive potential will be redistributed by climate change, as well as studying pollen as an aerosol in the climate system. While the empirical phenological models can be reproduced for any set of regional pollen counting stations, PECM as a whole can be easily adapted to various community climate and earth system models, global and regional, to extend research on the relationships and interactions between pollen and climate.

Matthew Wozniak 8/2/2017 7:28 PM
Deleted: estimating

7 Code and Data Availability

Source code for Pollen Emissions for Climate Models (PECM) is written as FORTRAN90 (*.f90) and available in the supplementary material as plain text. Input data is explained in Section 3 of this manuscript.

8 Acknowledgements

This work was supported by NSF Grant to AGS 0952659 to ALS. We thank Melissa Zagorski and Yang Li of University of Michigan for contributions to the emissions model development, and Fabien Solmon and Li Liu of the International Centre for Theoretical Physics for RegCM support and their prior work. We gratefully acknowledge the use of the American Academy of Allergy Asthma and Immunology (AAAAI) pollen count data, with individual station acknowledgments fully provided in Table S1.

Matthew Wozniak 8/2/2017 7:28 PM
Deleted: Association

Matthew Wozniak 8/2/2017 7:28 PM
Deleted: for

References

Beggs, Paul J, Branko Šikoparija, and Matt Smith. 2017. "Aerobiology in the International Journal of Biometeorology, 1957–2017." *International Journal of Biometeorology*, 1957–2017. Accessed August 8. doi:10.1007/s00484-017-1374-5.

Box, George E.P., Gwilym M. Jenkins, and Gregory C. Reinsel. 1994. *Time Series Analysis: Forecasting and Control*. Upper Saddle River, NJ: Prentice Hall. doi:10.1016/j.ijforecast.2004.02.001.

Cecchi, Lorenzo. 2014. "Introduction." In *Allergenic Pollen: A Review of the Production, Release, Distribution and Health Impacts*, edited by Mikhail Sofiev and Karl-Christian Bergmann, 1–7. New York, London: Springer Science+Business Media Dordrecht. doi:10.1007/978-94-007-4881-1.

Chuine, Isabelle, P. Cour, and D. D. Rousseau. 1999. "Selecting Models to Predict the Timing of Flowering of Temperate Trees: Implications for Tree Phenology Modelling." *Plant, Cell and Environment* 22 (1): 1–13. doi:10.1046/j.1365-3040.1999.00395.x.

Clay, S A, B Kreutner, D E Clay, C Reese, J Kleinjan, and F Forcella. 2006. "Spatial Distribution, Temporal Stability, and Yield Loss Estimates for Annual Grasses and Common Ragweed (Ambrosia Artimisiifolia) in a Corn/soybean Production Field over Nine Years." *Weed Science* 54 (2): 380–90. http://www.scopus.com/inward/record.url?eid=2-s2.0-33750295452&partnerID=40&md5=6931902680ebbd9303c2bd99f3ddc4a7.

Confalonieri, U., B. Menne, R. Akhtar, K.L. Ebi, M. Hauengue, R.S. Kovats, B. Revich, and A. Woodward. 2007. "Human Health." In *Climate Change 2007: Impacts, Adaptation and Vulnerability. Contribution of Working Group II to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*, edited by M.L. Parry, O.F. Canziani, J.P. Palutikof, P.J. van der Linden, and C.E. Hanson, 391–431. Cambridge, UK: Cambrige University Press.

Craine, Joseph M, Elizabeth M Wolkovich, E Gene Towne, and Steven W Kembel. 2011. "Flowering Phenology as a Functional Trait in a Tallgrass Prairie."

Dee, D P, S M Uppala, A J Simmons, P Berrisford, P Poli, S Kobayashi, U Andrae, et al. 2011. "The ERA-Interim Reanalysis: Configuration and Performance of the Data Assimilation System." *Quarterly Journal of the Royal Meteorological Society Q. J. R. Meteorol. Soc* 137: 553–97. doi:10.1002/qj.828.

Deen, William, Tony Hunt, Clarence J Swanton, and William Deen. 1998. "Influence of Temperature , Photoperiod , and Irradiance on the Phenological Development of Common Ragweed (Ambrosia Artemisiifolia)." *Weed Science* 46 (5): 555–60.

Despres, Viviane R., J. Alex Huffman, Susannah M. Burrows, Corinna Hoose, Aleksandr S. Safatov, Galina Buryak, Janine Frohlich-Nowoisky, et al. 2012. "Primary Biological Aerosol Particles in the Atmosphere: A Review." *Tellus, Series B: Chemical and Physical Meteorology*. doi:10.3402/tellusb.v64i0.15598.

Duhl, T. R., R. Zhang, a. Guenther, S. H. Chung, M. T. Salam, J. M. House, R. C. Flagan, et al. 2013. "The Simulator of the Timing and Magnitude of Pollen Season (STaMPS) Model: A Pollen Production Model for Regional Emission and Transport Modeling." *Geoscientific Model Development Discussions* 6 (2): 2325–68.

760 doi:10.5194/gmdd-6-2325-2013.
761 Easterling, David R. 2002. “Recent Changes in Frost Days and the Frost-Free Season in the United States.” *Bulletin*
762 *of the American Meteorological Society* 83: 1327–1332. http://journals.ametsoc.org/doi/pdf/10.1175/1520-
763 0477%282002%29083%3C1327%3ARCIFDA%3E2.3.CO%3B2.
764 Efstathiou, Christos, Sastry Isukapalli, and Panos Georgopoulos. 2011. “A Mechanistic Modeling System for
765 Estimating Large-Scale Emissions and Transport of Pollen and Co-Allergens.” *Atmospheric Environment* 45:
766 2260–76. doi:10.1016/j.atmosenv.2010.12.008.
767 Emanuel, Kerry. A. 1991. “A Scheme for Representing Cumulus Convention in Large-Scale Models.” *Journal of*
768 *the Atmospheric Sciences* 48 (21): 2313–35. ftp://texmex.mit.edu/pub/emanuel/PAPERS/convect91.pdf.
769 Emberlin, J., J. Mullins, J. Corden, S. Jones, W. Millington, M. Brooke, and M. Savage. 1999. “Regional Variations
770 in Grass Pollen Seasons in the UK, Long-Term Trends and Forecast Models.” *Clinical and Experimental*
771 *Allergy* 29 (3): 347–56. doi:10.1046/j.1365-2222.1999.00369.x.
772 Forman, Richard T T, and Lauren E Alexander. 1998. “ROADS AND THEIR MAJOR ECOLOGICAL EFFECTS.”
773 *Annu. Rev. Ecol. Syst* 29: 207–31.
774 Fu, Yongshuo H., Matteo Campioli, Gaby Deckmyn, and Ivan A. Janssens. 2012. “The Impact of Winter and Spring
775 Temperatures on Temperate Tree Budburst Dates: Results from an Experimental Climate Manipulation.”
776 *PLoS ONE* 7 (10). doi:10.1371/journal.pone.0047324.
777 Fumanal, Boris, Bruno Chauvel, and Francois Bretagnolle. 2007. “ESTIMATION OF POLLEN AND SEED
778 PRODUCTION OF COMMON RAGWEED IN FRANCE.” *Ann Agric Environ Med* 14: 233–36.
779 doi:10.1093/annonc/mdw163.
780 Galán, C, García-, H Mozo, L Vázquez, L Ruiz, C Guardia, and Díaz. 2008. “Modeling Olive Crop Yield in
781 Andalusia, Spain.” *E Agronomy Journal* 100 (1).
782 http://search.proquest.com/docview/194514400/fulltextPDF/5D1BBD9891A04649PQ/1?accountid=14667.
783 García-Mozo, H., C. Galán, J. Belmonte, D. Bermejo, P. Candau, C. Díaz de la Guardia, B. Elvira, et al. 2009.
784 “Predicting the Start and Peak Dates of the Poaceae Pollen Season in Spain Using Process-Based Models.”
785 *Agricultural and Forest Meteorology* 149 (2): 256–62. doi:10.1016/j.agrformet.2008.08.013.
786 Georg a, Grell, Jimy Dudhia, and David R Stauffer. 1994. “A Description of the Fifth-Generation Penn State/NCAR
787 Mesoscale Model (MM5).” *NCAR Technical Note NCAR/TN-398+STR*, no. December: 121.
788 doi:10.5065/D60Z716B.
789 Giorgi, F., E. Coppola, F. Solmon, L. Mariotti, M. B. Sylla, X. Bi, N. Elguindi, et al. 2012. “RegCM4: Model
790 Description and Preliminary Tests over Multiple CORDEX Domains.” *Climate Research* 52 (1): 7–29.
791 doi:10.3354/cr01018.
792 Guenther, A., T. Karl, P. Harley, C. Wiedinmyer, P. I. Palmer, and C. Geron. 2006. “Estimates of Global Terrestrial
793 Isoprene Emissions Using MEGAN (Model of Emissions of Gases and Aerosols from Nature).” *Atmospheric*
794 *Chemistry and Physics* 6: 3181–3210. doi:10.1016/j.cognition.2008.05.007.
795 Harris, I, P D Jones, T J Osborn, and D H Lister. 2014. “Updated High-Resolution Grids of Monthly Climatic
796 Observations – the CRU TS3 . 10 Dataset” 642 (May 2013): 623–42. doi:10.1002/joc.3711.

Helbig, Nora, Bernhard Vogel, Heike Vogel, and Franz Fiedler. 2004. “Numerical Modelling of Pollen Dispersion on the Regional Scale.” *Aerobiologia* 3: 3–19.

Hidalgo, Pablo J, Carmen Galán, and Eugenio Domínguez. 1999. “Pollen Production of the Genus Cupressus.” *Grana* 38 (5): 296–300. doi:10.1080/001731300750044519.

Hidalgo, Pablo J, Antoine Mangin, Carmen Galán, Odile Hembise, Luis M Vázquez, and Oscar Sanchez. 2002. “An Automated System for Surveying and Forecasting Olea Pollen Dispersion.” *Aerobiologia* 18: 23–31.

Hunt, J. C. R., H. L. Higson, P. J. Walklate, and J. B. Sweet. 2002. “Modelling the Dispersion and Cross-Fertilisation of Pollen from GM Crops.”

Jackson, Trisha L, Johannes J Feddema, Keith W Oleson, Gordon B Bonan, and John T Bauer. 2010. “Parameterization of Urban Characteristics for Global Climate Modeling.” doi:10.1080/00045608.2010.497328.

Jato, Victoria, F. Javier Rodríguez-Rajo, and M. Jesús Aira. 2007. “Use of Phenological and Pollen-Production Data for Interpreting Atmospheric Birch Pollen Curves.” *Annals of Agricultural and Environmental Medicine* 14 (2): 271–80.

Jochner, Susanne, Josef Hofler, Isabelle Beck, Axel Gottlein, Donna Pauler Ankerst, Claudia Traidl-Hoffmann, and Annette Menzel. 2013. “Nutrient Status: A Missing Factor in Phenological and Pollen Research?” *Journal of Experimental Botany* 64 (7): 2081–92. doi:10.1093/jxb/ert061.

Katz, Daniel S W, Benjamin T. Connor Barrie, and Tiffany S. Carey. 2014. “Urban Ragweed Populations in Vacant Lots: An Ecological Perspective on Management.” *Urban Forestry and Urban Greening* 13 (4). Elsevier GmbH.: 756–60. doi:10.1016/j.ufug.2014.06.001.

Kinnee, Ellen, Chris Geron, and Thomas Pierce. 1997. “United States Land Use Inventory For Estimating Biogenic Ozone Precursor Emissions.” *Ecological Applications* 7 (1): 46.

Kuparinen, Anna, Tiina Markkanen, Hermanni Riikonen, and Timo Vesala. 2007. “Modeling Air-Mediated Dispersal of Spores, Pollen and Seeds in Forested Areas.” doi:10.1016/j.ecolmodel.2007.05.023.

Lake, Iain R., Natalia R. Jones, Maureen Agnew, Clare M. Goodess, Filippo Giorgi, Lynda Hamaoui-Laguel, Mikhail A. Semenov, et al. 2017. “Climate Change and Future Pollen Allergy in Europe.” *Environmental Health Perspectives* 125 (3): 385–91. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5332176/pdf/EHP173.pdf.

Larson, Diane L. 2003. “Native Weeds and Exotic Plants: Relationships to Disturbance in Mixed-Grass Prairie.” *Plant Ecology* 169 (2): 317–33. doi:10.1023/A:1026046810307.

Lawrence, Peter J., and Thomas N. Chase. 2007. “Representing a New MODIS Consistent Land Surface in the Community Land Model (CLM 3.0).” *Journal of Geophysical Research: Biogeosciences* 112 (1). doi:10.1029/2006JG000168.

Lewis, Walter H., Prathibha. Vinay, and Vincent E. Zenger. 1983. *Airborne and Allergenic Pollen of North America*. Baltimore: Johns Hopkins University Press. https://mirlyn.lib.umich.edu/Record/000779939.

Linkosalo, Tapio, Hanna K Lappalainen, and Pertti Hari. 2008. “A Comparison of Phenological Models of Leaf Bud Burst and Flowering of Boreal Trees Using Independent Observations.” *Tree Physiology* 28 (12): 1873–82.

834 doi:10.1093/treephys/28.12.1873.
835 Liu, Li, Fabien Solmon, Robert Vautard, Lynda Hamaoui-laguel, Csaba Zsolt Torma, and Filippo Giorgi. 2016.
836 “Ragweed Pollen Production and Dispersion Modelling within a Regional Climate System , Calibration and
837 Application over Europe,” 2769–86. doi:10.5194/bg-13-2769-2016.
838 Maurer, E. P., A. W. Wood, J. C. Adam, D. P. Lettenmaier, and B. Nijssen. 2002. “A Long-Term Hydrologically
839 Based Dataset of Land Surface Fluxes and States for the Conterminous United States: Update and
840 Extensions.” *Journal of Climate*, 3237–51. doi:http://dx.doi.org/10.1175/1520-
841 0442(2002)015<3237:ALTHBD>2.0.CO;2.
842 Medek, Danielle E, Paul J. Beggs, Bircan Erbas, Alison K. Jaggard, Bradley C. Campbell, Don Vicendese, Fay H
843 Johnston, et al. 2016. “Regional and Seasonal Variation in Airborne Grass Pollen Levels between Cities of
844 Australia and New Zealand.” *Aerobiologia* 32 (2): 289–302. doi:10.1007/s10453-015-9399-x.
845 Menzel, Annette, and Susanne Jochner. 2016. “Impacts of Climate Change on Aeroallergen Production and
846 Atmospheric Concentration.” In *Impacts of Climate Change on Allergens and Allergic Diseases*, edited by
847 Paul J. Beggs, 10–28. Cambridge, UK: Cambridge University Press.
848 Moseholm, L., E. R. Weeke, and B. N. Petersen. 1987. “Forecast of Pollen Concentrations of Poaceae (Grasses) in
849 the Air by Time Series Analysis.” *Pollen et Spores* 29 (2–3): 305–21.
850 Myking, T., and O. M. Heide. 1995. “Dormancy Release and Chilling Requirement of Buds of Latitudinal Ecotypes
851 of Betula Pendula and B. Pubescens.” *Tree Physiology* 15 (11). Oxford University Press: 697–704.
852 doi:10.1093/treephys/15.11.697.
853 Myriokefalitakis, S, G Fanourgakis, and M Kanakidou. 2017. “The Contribution of Bioaerosols to the Organic
854 Carbon Budget of the Atmosphere.” In *Perspectives on Atmospheric Sciences*, 845–51. doi:10.1007/978-3-
855 319-35095-0_121.
856 Oleson, K, D Lawrence, G Bonan, M Flanner, and Kluzek E. 2010. “Technical Description of Version 4.0 of the
857 Community Land Model (CLM).” *NCAR Technical Note NCAR/TN-478+STR*, 257.
858 Olsson, Cecilia, and Anna Maria Jönsson. 2014. “Process-Based Models Not Always Better than Empirical Models
859 for Simulating Budburst of Norway Spruce and Birch in Europe.” *Global Change Biology*, 3492–3507.
860 doi:10.1111/gcb.12593.
861 Pal, Jeremy S, Eric E Small, and Elfatih A B Eltahir. 2000. “Simulation of Regional-Scale Water and Energy
862 Budgets: Representation of Subgrid Cloud and Precipitation Processes within RegCM.” *JOURNAL OF
863 GEOPHYSICAL RESEARCH* 105594 (27): 579–29. http://geode.colorado.edu/~small/docs/2000JGR.pdf.
864 Prieto-Baena, José C, Pablo J Hidalgo, Carmen Galán, and Eugenio Domínguez. 2003. “Pollen Production in the
865 Poaceae Family.” *Grana* 42 (3): 153–59.
866 http://www.tandfonline.com/doi/pdf/10.1080/00173130310011810?needAccess=true.
867 Reichstein, Markus, Michael Bahn, Miguel D. Mahecha, Jens Kattge, and Dennis D. Baldocchi. 2014. “Linking
868 Plant and Ecosystem Functional Biogeography.” *Proceedings of the National Academy of Sciences* 111 (38):
869 201216065. doi:10.1073/pnas.1216065111.
870 Scheifinger, Helfried, Jordina Belmonte, Jeroen Buters, Sevcan Celenk, Athanasios Damialis, Chantal Dechamp,

Herminia García-Mozo, et al. 2013. “Monitoring, Modelling and Forecasting of the Pollen Season.” In *Allergenic Pollen*, edited by Mikhail Sofiev and Karl-Christian Bergmann. New York, London: Springer Science+Business Media Dordrecht. doi:10.1007/978-94-007-4881-1_4.

Schueler, Silvio, and Katharina Heinke Schlünzen. 2006. “Modeling of Oak Pollen Dispersal on the Landscape Level with a Mesoscale Atmospheric Model.” *Environmental Modeling and Assessment* 11 (3): 179–94. doi:10.1007/s10666-006-9044-8.

Siljamo, Pilvi, Mikhail Sofiev, Elena Filatova, Łukasz Grewling, Siegfried Jäger, Ekaterina Khoreva, Tapio Linkosalo, et al. 2013. “A Numerical Model of Birch Pollen Emission and Dispersion in the Atmosphere. Model Evaluation and Sensitivity Analysis.” *International Journal of Biometeorology* 57 (1): 125–36. doi:10.1007/s00484-012-0539-5.

Smith, M., and J. Emberlin. 2005. “Constructing a 7-Day Ahead Forecast Model for Grass Pollen at North London, United Kingdom.” *Clinical and Experimental Allergy* 35 (10): 1400–1406. doi:10.1111/j.1365-2222.2005.02349.x.

Smith, Thomas M, Richard W Reynolds, Thomas C Peterson, and Jay Lawrimore. 2008. “Improvements to NOAA’s Historical Merged Land-Ocean Surface Temperature Analysis (1880-2006).” *Journal of Climate* 21 (10): 2283–96. doi:10.1175/2007JCLI2100.1.

Sofiev, M., P. Siljamo, H. Ranta, T. Linkosalo, S. Jaeger, A. Rasmussen, A. Rantio-Lehtimaki, E. Severova, and J. Kukkonen. 2013. “A Numerical Model of Birch Pollen Emission and Dispersion in the Atmosphere. Description of the Emission Module.” *International Journal of Biometeorology* 57 (1): 45–58. doi:10.1007/s00484-012-0532-z.

Sofiev, M., P. Siljamo, H. Ranta, and A. Rantio-Lehtimäki. 2006. “Towards Numerical Forecasting of Long-Range Air Transport of Birch Pollen: Theoretical Considerations and a Feasibility Study.” *International Journal of Biometeorology* 50 (6): 392–402. doi:10.1007/s00484-006-0027-x.

Sofiev, Mikhail, Jordina Belmonte, Regula Gehrig, Rebeca Izquierdo, Matt Smith, Aslog Dahl, and Pilvi Siljamo. 2014. “Airborne Pollen Transport.” In *Allergenic Pollen: A Review of the Production, Release, Distribution and Health Impacts*, edited by Mikhail Sofiev and Karl-Christian Bergman, 127–59. New York, London: Springer Science+Business Media Dordrecht. doi:10.1007/978-94-007-4881-1.

Sofiev, Mikhail, and Marje Prank. 2016. “Impacts of Climate Change on Aeroallergen Dispersion, Transport, and Deposition.” In *Impacts of Climate Change on Allergens and Allergic Diseases*, edited by Paul J. Beggs, 50–73. Cambridge, UK: Cambrige University Press.

Solmon, F., F. Giorgi, and C. Liousse. 2006. “Aerosol Modelling for Regional Climate Studies: Application to Anthropogenic Particles and Evaluation over a European/African Domain.” *Tellus, Series B: Chemical and Physical Meteorology* 58 (1): 51–72. doi:10.1111/j.1600-0889.2005.00155.x.

Thornton, P.E., M.M. Thornton, B.W. Mayer, Nate Wilhelmi, Yaxing Wei, Ranjeet Devarakonda, and R.B. Cook. 2014. “Daymet: Daily Surface Weather Data on a 1-Km Grid for North America, Version 2. Data Set.” *Oak Ridge National Laboratory Distributed Active Archive Center, Oak Ridge, Tennessee, USA.* doi:http://dx.doi.org/10.3334/ORNLDAAC/1219.

Tormo Molina, Rafael, Adolfo Muñoz Rodríguez, Silva Palacios, and Francisco Gallardo Lopes. 1996. "Pollen Production in Anemophilous Trees." *Grana* 35: 38–46. doi:10.1080/00173139609430499.

Weber, Richard W. 2003. "Meteorologic Variables in Aerobiology." *Immunology and Allergy Clinics of North America* 23 (3): 411–22. doi:10.1016/S0889-8561(03)00062-6.

Yue, X., N. Unger, T. F. Keenan, X. Zhang, and C. S. Vogel. 2015. "Probing the Past 30-Year Phenology Trend of US Deciduous Forests." *Biogeosciences* 12 (15): 4693–4709. doi:10.5194/bg-12-4693-2015.

Zhang, R., T. Duhl, M. T. Salam, J. M. House, R. C. Flagan, E. L. Avol, F. D. Gilliland, et al. 2014. "Development of a Regional-Scale Pollen Emission and Transport Modeling Framework for Investigating the Impact of Climate Change on Allergic Airway Disease." *Biogeosciences* 11 (6): 1461–78. doi:10.5194/bg-11-1461-2014.

Zhang, Yong, Leonard Bielory, Ting Cai, Zhongyuan Mi, and Panos Georgopoulos. 2015. "Predicting Onset and Duration of Airborne Allergenic Pollen Season in the United States." *Atmospheric Environment* 103. Elsevier Ltd: 297–306. doi:10.1016/j.atmosenv.2014.12.019.

Zhang, Yong, Leonard Bielory, Zhongyuan Mi, Ting Cai, Alan Robock, and Panos Georgopoulos. 2015. "Allergenic Pollen Season Variations in the Past Two Decades under Changing Climate in the United States." *Global Change Biology* 21 (4): 1581–89. doi:10.1111/gcb.12755.

Ziello, Chiara, Tim H Sparks, Nicole Estrella, Jordina Belmonte, Karl C Bergmann, Edith Bucher, Maria Antonia Brighetti, et al. 2012. "Changes to Airborne Pollen Counts across Europe." *PLoS ONE* 7 (4). doi:10.1371/journal.pone.0034076.

Zink, K., a. Pauling, M. W. Rotach, H. Vogel, P. Kaufmann, and B. Clot. 2013. "EMPOL 1.0: A New Parameterization of Pollen Emission in Numerical Weather Prediction Models." *Geoscientific Model Development* 6 (6): 1961–75. doi:10.5194/gmd-6-1961-2013.

Zink, Katrin, Pirmin Kaufmann, Blaise Petitpierre, Olivier Broennimann, Antoine Guisan, Eros Gentilini, Mathias W Rotach, Mathias W Rotach MathiasRotach, and uibkacat Katrin Zink. 2017. "Numerical Ragweed Pollen Forecasts Using Different Source Maps: A Comparison for France." *Int J Biometeorol* 61: 23–33. doi:10.1007/s00484-016-1188-x.

Ziska, Lewis H. 2016. "Impacts of Climate Change on Allergen Seasonality." In *Impacts of Climate Change on Allergens and Allergic Diseases*, edited by Paul J. Beggs, 93–112. Cambridge, UK: Cambrige University Press.

Ziska, Lewis, Kim Knowlton, Christine Rogers, Dan Dalan, Nicole Tierney, Mary Ann Elder, Warren Filley, et al. 2011. "Recent Warming by Latitude Associated with Increased Length of Ragweed Pollen Season in Central North America." *Proceedings of the National Academy of Sciences of the United States of America* 108 (10): 4248–51. doi:10.1073/pnas.1014107108.

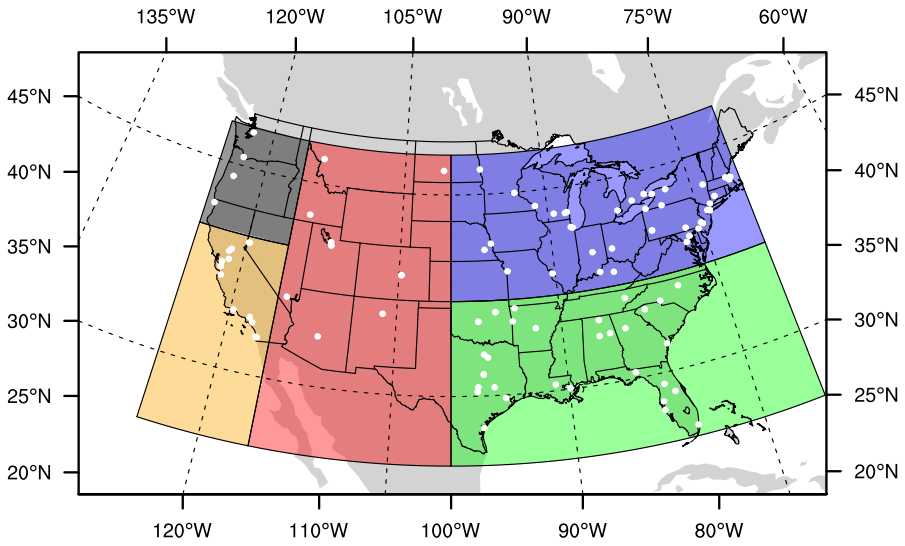


Figure 1. Locations of AAAAI station locations and geographic regions used in this study: Northeast (NE; 38°-48°N, 70°-100°W) in blue, Southeast (SE; 25°-38°N, 70°-100°W) in green, Mountain (MT; 25°-48°N, 100°-116°W) in red, California (CA; 25°-40°N, 116°-125°W) in orange, and Pacific Northwest (PNW; 40°-48°N, 116°-125°W) in dark grey.

Matthew Wozniak 8/9/2017 10:06 PM
Deleted: >

Matthew Wozniak 8/9/2017 10:06 PM
Deleted: <

Matthew Wozniak 8/9/2017 10:06 PM
Deleted: <

Matthew Wozniak 8/9/2017 10:07 PM
Deleted: <

Matthew Wozniak 8/9/2017 10:08 PM
Deleted: W to

Matthew Wozniak 8/9/2017 10:08 PM
Deleted: <

Matthew Wozniak 8/9/2017 10:08 PM
Deleted: >

Matthew Wozniak 8/9/2017 10:08 PM
Deleted:

Matthew Wozniak 8/9/2017 10:09 PM
Deleted: >

Matthew Wozniak 8/9/2017 10:09 PM
Deleted: >

Matthew Wozniak 8/9/2017 10:09 PM
Deleted:

Matthew Wozniak 8/2/2017 7:29 PM
Deleted: black

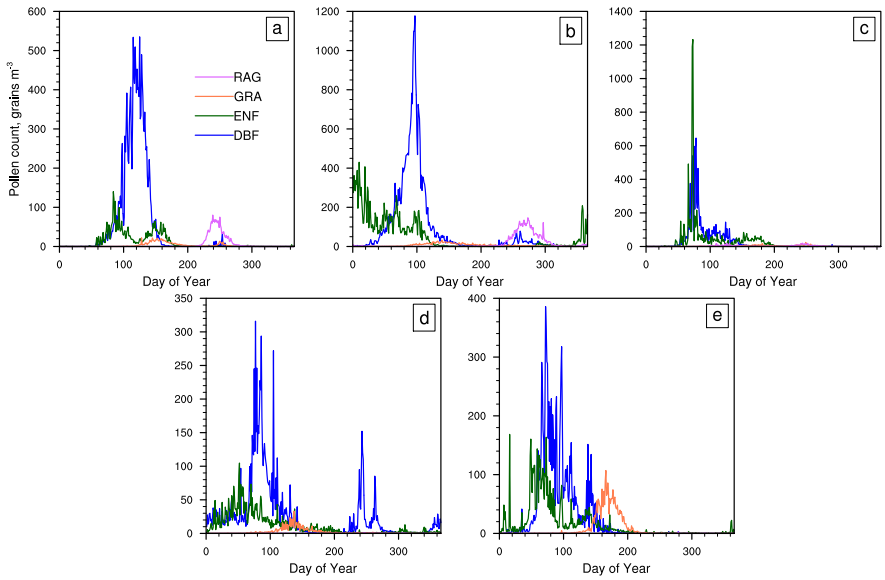


Figure 2. Daily observed average time series of daily pollen count data (2003-2010) for the four representative plant functional types (DBF, ENF, grasses, ragweed) averaged over the five regions in Figure 1: (a) Northeast, (b) Southeast, (c) Mountain, (d) California, and (e) Pacific Northwest.

Unknown
Formatted: Font:Bold

Matthew Wozniak 4/28/2017 3:04 PM
Comment [1]: Up the font in axes.

Matthew Wozniak 8/8/2017 11:01 PM
Deleted: <*sp*>

Matthew Wozniak 8/2/2017 7:29 PM
Deleted: climatological

Unknown
Formatted: Font:Bold

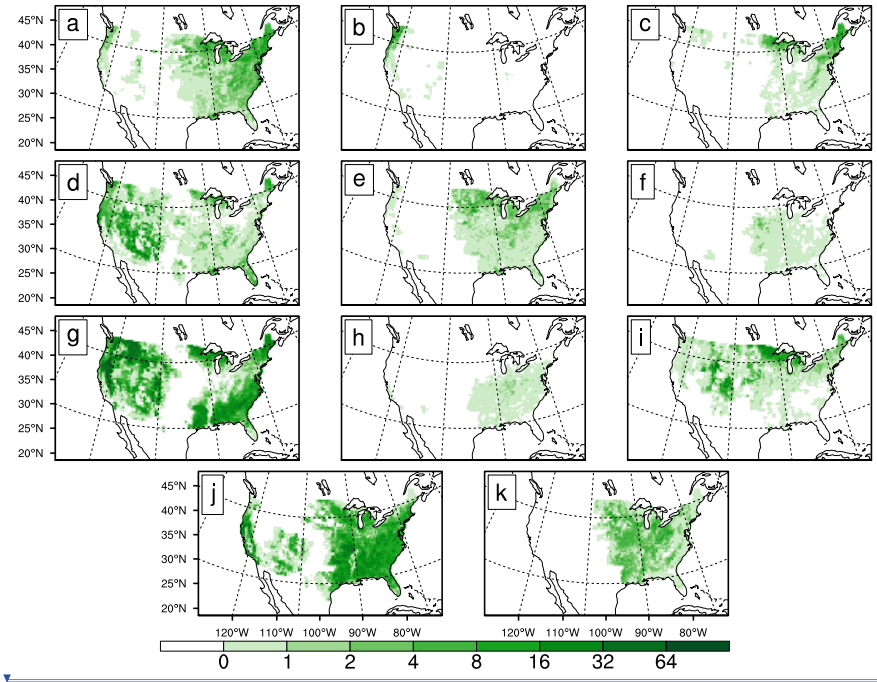
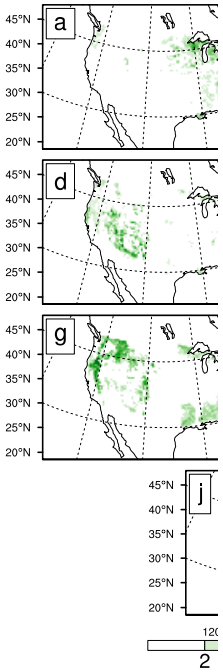


Figure 3. Land cover fraction (% coverage) for 11 tree taxa from the Biogenic Emissions Landuse Database (BELD3) regridded to a 25km resolution grid, including: a) *Acer* (maple), b) *Alnus* (alder), c) *Betula* (birch), d) Cupressaceae (cedar/juniper), e) *Fraxinus* (ash), f) *Morus* (mulberry), g) Pinaceae (pine), h) *Platanus* (sycamore), i) *Populus* (poplar/aspen), j) *Quercus* (oak), k) *Ulmus* (elm).

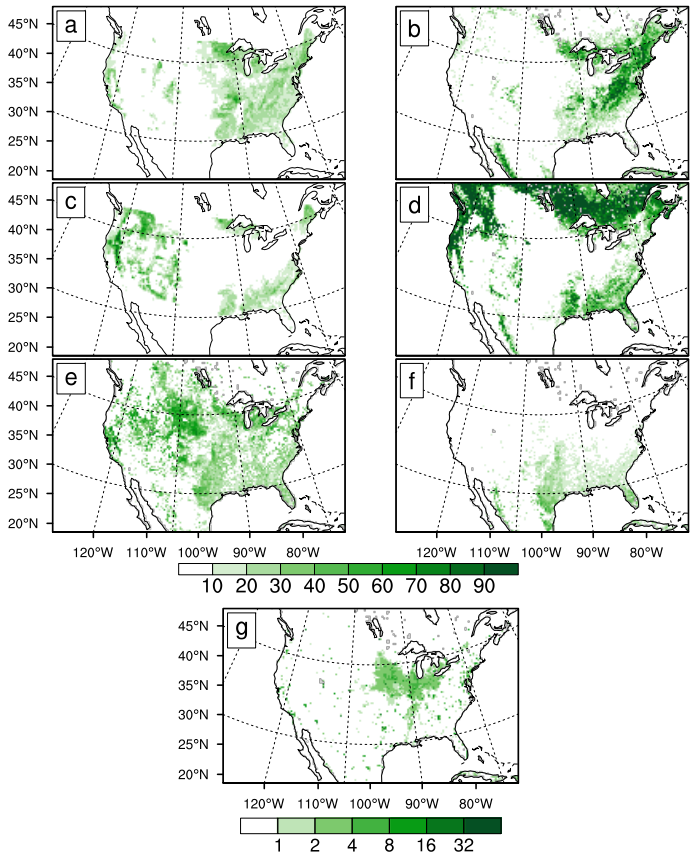
Matthew Wozniak 8/10/2017 2:57 PM



Deleted:

Matthew Wozniak 8/2/2017 7:30 PM

Deleted: Land cover



978

979 **Figure 4.** BELD3 (a, c) and CLM4 (b, d, e, f) land cover for the four PFT categories that produce pollen emissions,
980 including (1) deciduous broadleaf forest for (a) BELD3 and (b) CLM4, (2) evergreen needleaf forest for (c) BELD3
981 and (d) CLM4, (3) grasses, including (e) C3 grasses and (f) C4 grasses, and (g) ragweed, represented by crop and
982 urban CLM4 categories.

983

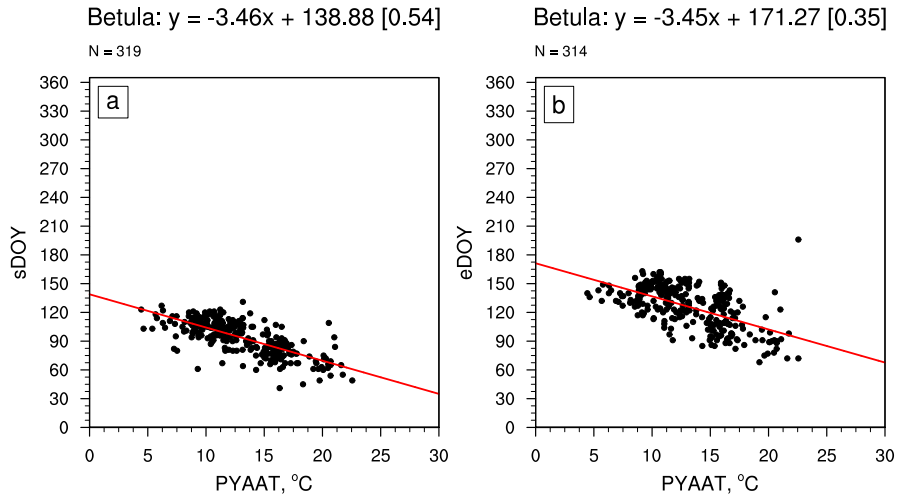
Matthew Wozniak 8/10/2017 3:06 PM
Deleted: IGBP

Matthew Wozniak 8/10/2017 3:06 PM
Deleted: IGBP

Matthew Wozniak 8/10/2017 3:06 PM
Deleted: IGBP

Matthew Wozniak 8/2/2017 7:31 PM
Deleted: 4

Matthew Wozniak 8/10/2017 3:06 PM
Deleted: IGBP



989

990 **Figure 5.** Phenological regressions for *Betula* (birch) pollen for (a) Start Day of Year (sDOY) and (b) End Day of
991 Year (eDOY) versus previous year annual average temperature (PYAAT; ºC). Each point signifies one station per
992 year for pollen count data from 2003-2010 (total denoted as N).

993

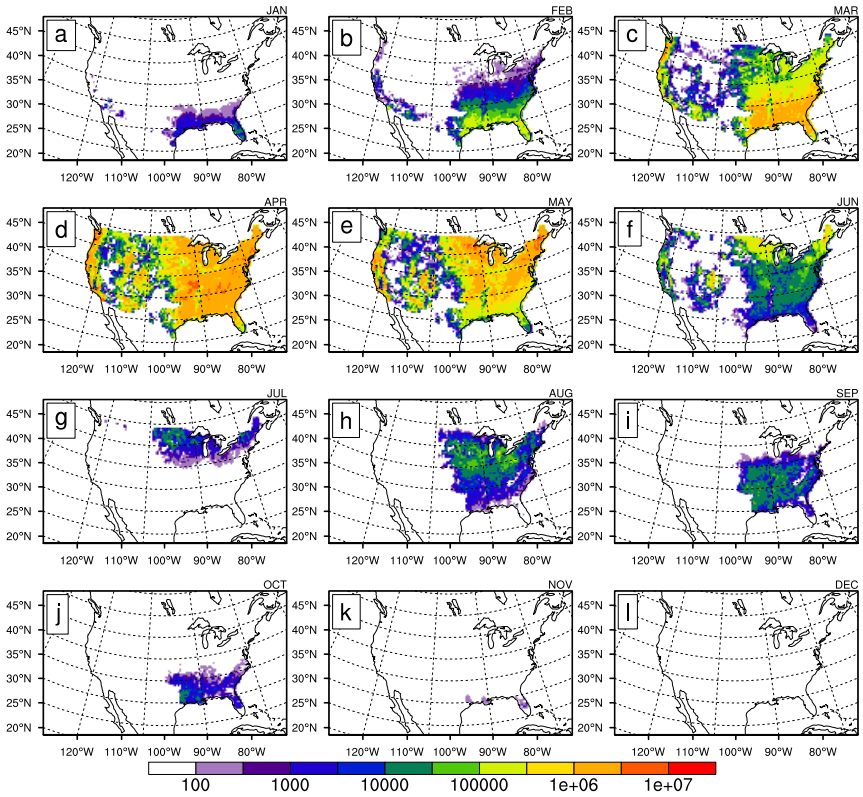


Figure 6. Monthly average emissions potential (E; Equation 1) for BELD model DBF (2003-2010), in grains m^{-2} day^{-1}. a) January, b) February, c) March, d) April, e) May, f) June, g) July, h) August, i) September, j) October, k) November, l) December.

Matthew Wozniak 8/2/2017 7:32 PM
Deleted: climatological

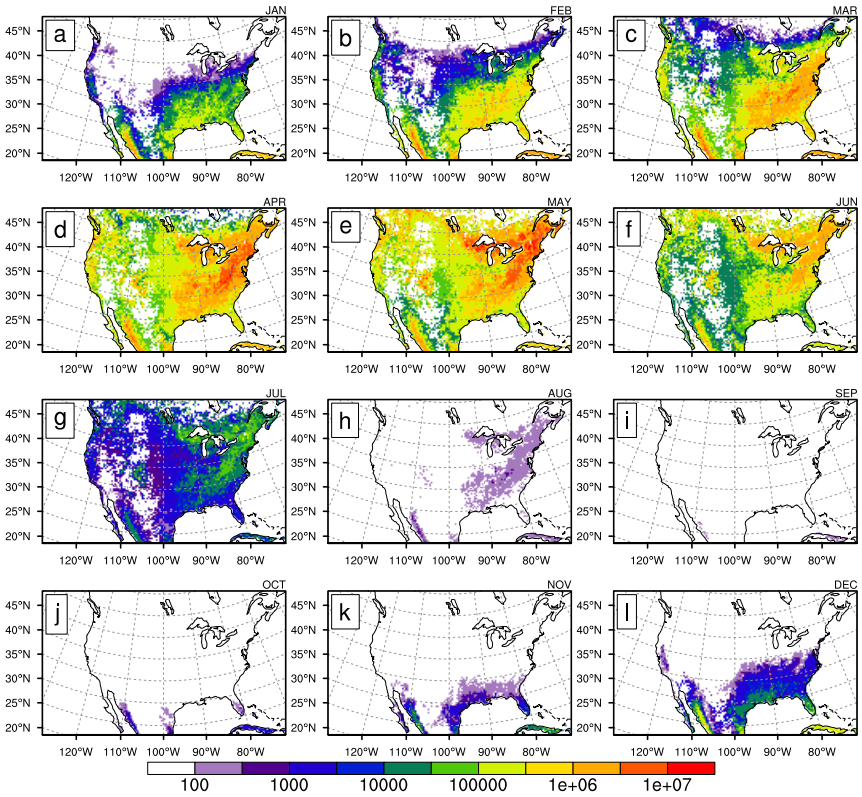


Figure 7. Same as Figure 6, but for PFT model DBF.

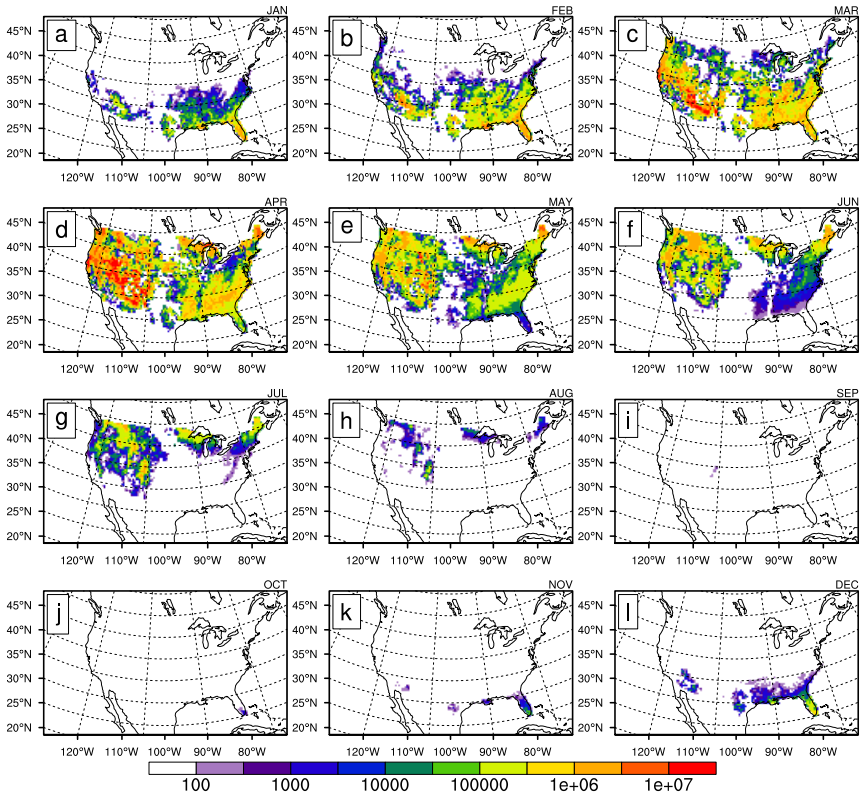


Figure 8. Same as Figure 6, but for BELD model ENF.

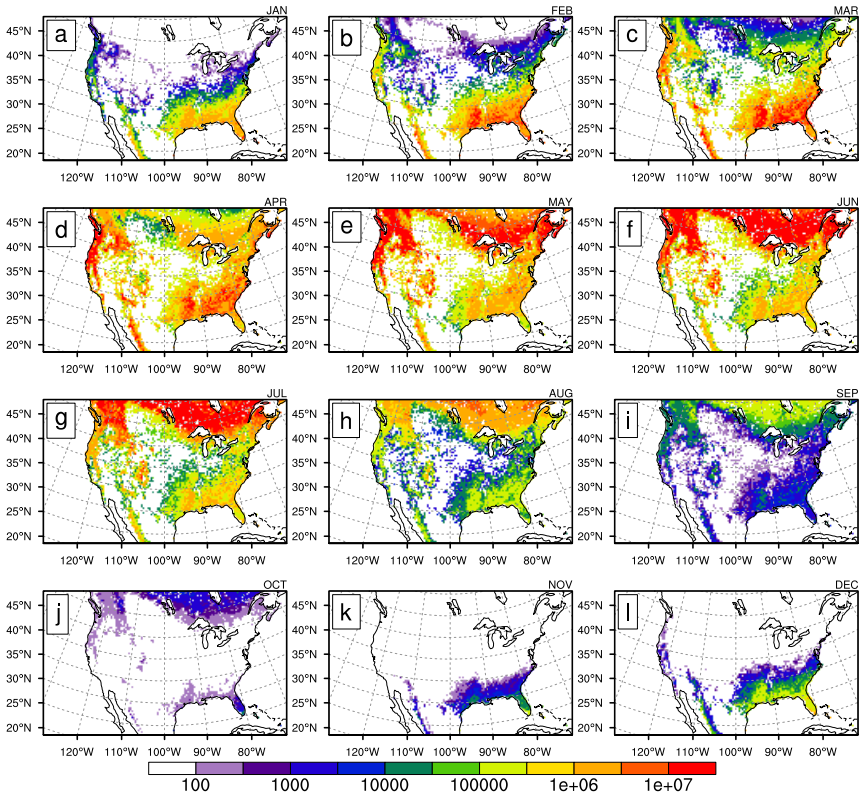


Figure 9. Same as Figure 6, but for PFT model ENF.

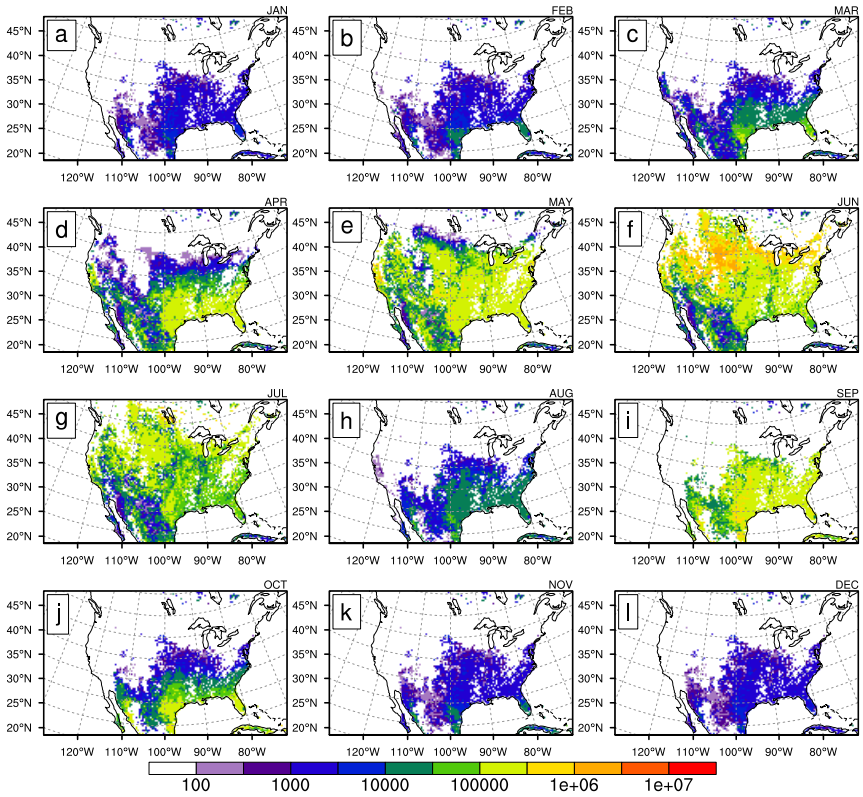


Figure 10. Same as Figure 6, but for C_3 + C_4 grass.

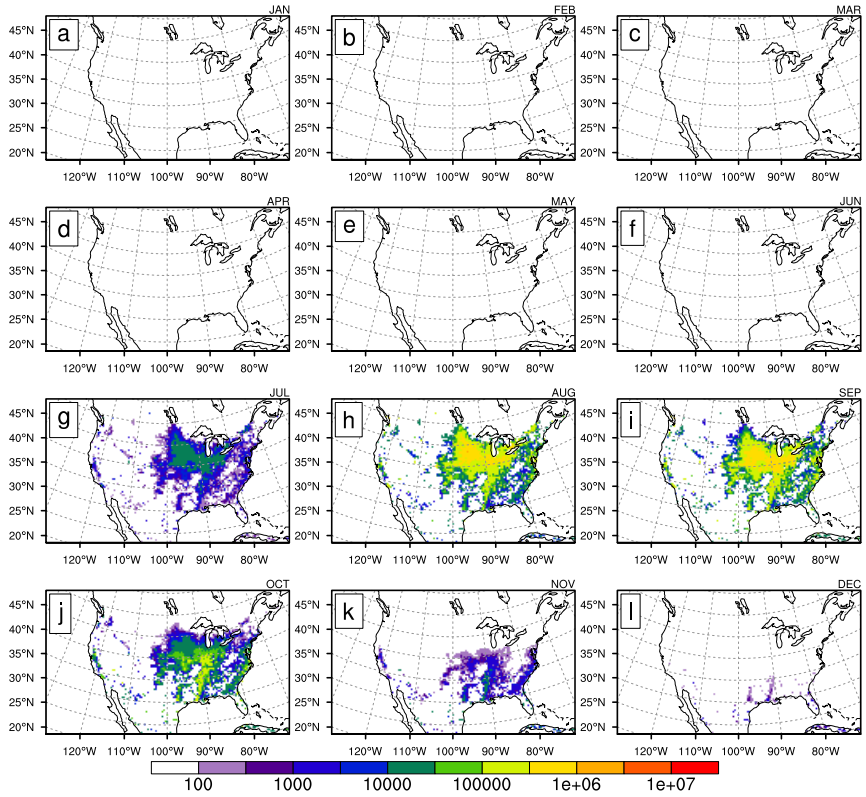


Figure 11. Same as Figure 6, but for ragweed.

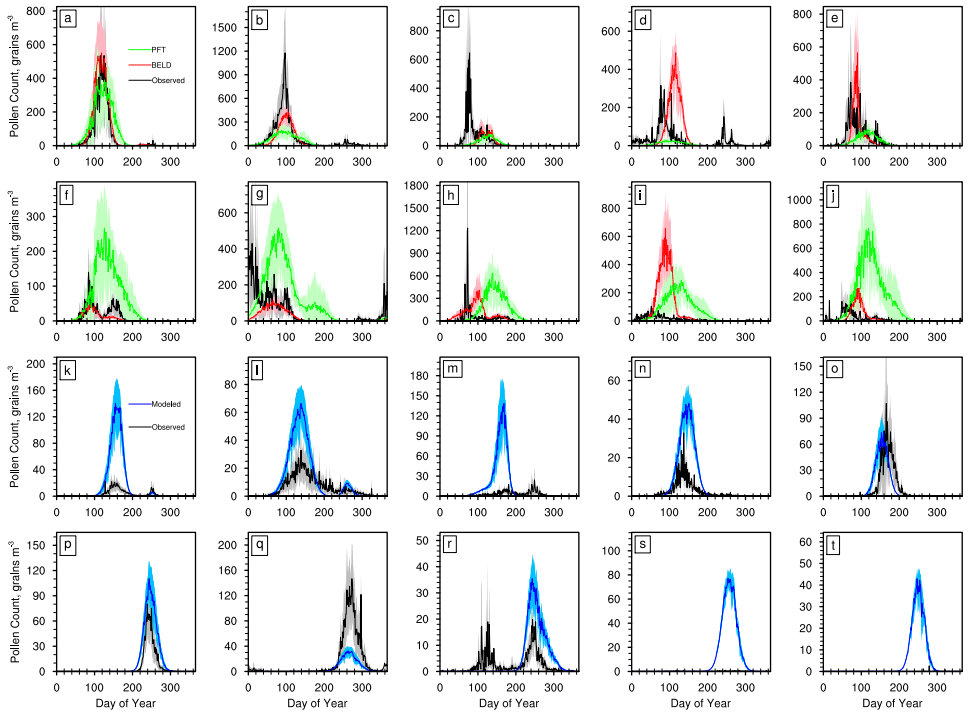


Figure 12. Average (2003-2010) time series of daily pollen counts comparing model and observations for four PFTs (a-e, deciduous broadleaf, DBF; f-j, evergreen needleleaf, ENF; k-o, grasses, GRA; p-t, ragweed, RAG) across 5 U.S. subregions (columns from left to right: Northeast, NE; Southeast, SE; Mountain, MT; California; Pacific Northwest, PNW). Shading for the observations and model represents the mean absolute deviation from the average for each day of the time series. Note: scale of y-axes varies by region and PFT.

Matthew Wozniak 8/8/2017 10:26 PM

Deleted:

Unknown

Formatted: Font:Bold

Matthew Wozniak 8/2/2017 7:32 PM

Deleted: Daily climatological

Matthew Wozniak 8/2/2017 7:33 PM

Deleted: u

Matthew Wozniak 8/2/2017 7:32 PM

Deleted: climatological

Matthew Wozniak 8/2/2017 7:33 PM

Deleted: climatology

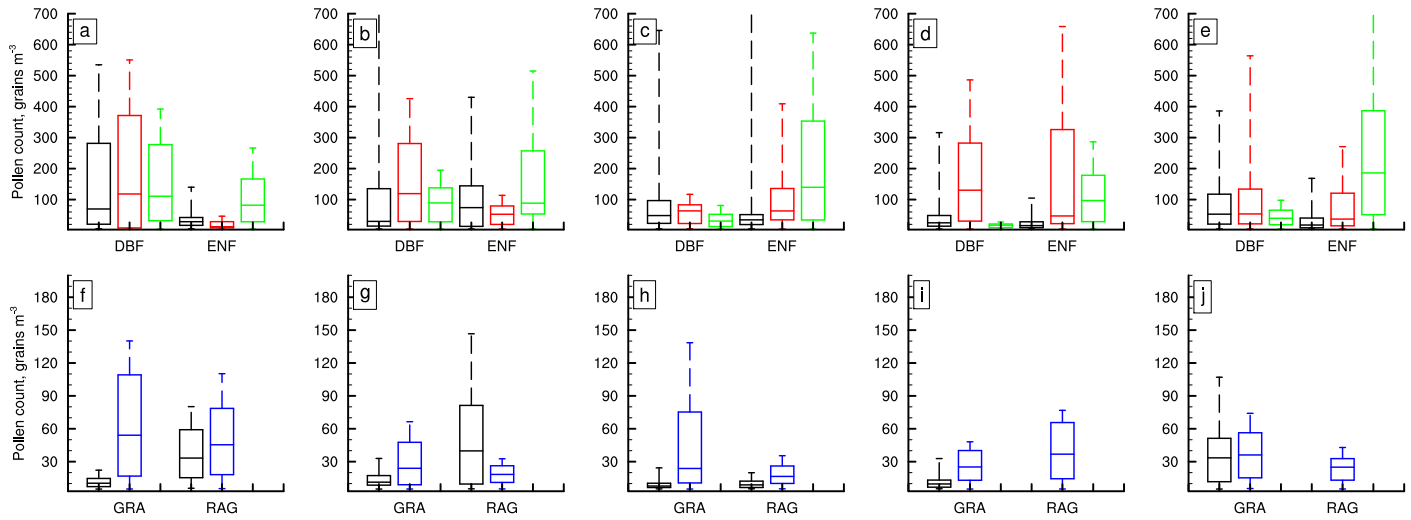


Figure 13. Box-and-whisker plots showing the statistical spread of the pollen count magnitudes from the regional averages presented in Figure 12. Columns from left to right: Northeast, NE (a,f); Southeast, SE (b, g) ; Mountain, MT (c,h); California (d,i); Pacific Northwest, PNW (e,j). DBF and ENF PFTs are shown in the top row (a-e) and grass and ragweed PFTs are shown in the bottom row (f-j). Box and whiskers from bottom to top represent the minimum, lower quartile, median, upper quartile, and maximum. Maxima that are not visible in panels b, c and e are 1,177 grains m^{-3}, 1,233 grains m^{-3}, and 766 gains m^{-3} respectively. All y-axes are the same scale for each row.

Matthew Wozniak 8/17/2017 8:51 PM
Deleted:

Matthew Wozniak 8/17/2017 8:51 PM
Deleted: noted on plot as a number

Taxon or PFT	P 10^7 grains m^{-2} $year^{-1}$	Reference for P	sDOY (slope/R^2) $days\ ºC^{-1}$	eDOY (slope/R^2) $days\ ºC^{-1}$
Deciduous Broadleaf Forest (DBF)				
Acer	89.1	Tormo Molina et al. 1996	-1.78/0.15	-1.56/0.06
Alnus	210	Helbig et al. 2004	-8.82/0.46	-4.88/0.26
Betula	140	Jato et al. 2007	-3.46/0.54	-3.45/0.35
Fraxinus	45.1	Tormo Molina et al. 1996	-4.69/0.50	-2.92/0.32
Morus	10	N/A	-4.00/0.53	-2.97/0.29
Platanus	121	Tormo Molina et al. 1996	-4.47/0.40	-2.65/0.20
Populus	24.2	Tormo Molina et al. 1996	-2.23/0.24	-0.31/<0.01
Quercus	78	Tormo Molina et al. 1996	-4.09/0.53	-2.03/0.19
Ulmus (early,late)	3.55	Tormo Molina et al. 1996	-4.61/0.59, 3.06/0.12	-2.37/0.16, 5.12/0.29
DBF	80.1		-4.55/0.46	-1.94/0.13
Evergreen Needleleaf Forest (ENF)				
Cupressaceae	363	Hidalgo et al. 1999	-5.67/0.48	-2.67/0.17
Pinaceae	22.2	Tormo Molina et al. 1996	-5.72/0.45	-5.03/0.41
ENF	193		-5.95/0.40	-4.96/0.33
Grasses (GRA)				
Poaceae (C_3,C_4)	8.5, 0.85	Prieto-Baena et al. 2003	-4.76/0.48, 0.05/<0.01	-1.08/0.04, 2.96/0.32
Ragweed (RAG)				
Ambrosia	119 [a]	Fumanal et al. 2007	1.08/0.08	3.42/0.37

[a] *Ambrosia* production factor in 10^7 grains $plant^{-1}$

Table 1: Production factors (P) and phenological regression coefficients for the start day of year (sDOY) and end day of year (eDOY) as a function of temperature for the 13 individual pollen-producing taxa. Individual taxa and

Matthew Wozniak 8/19/2017 6:33 AM
Deleted: (

Matthew Wozniak 8/19/2017 6:33 AM
Deleted:)

Allison Steiner 8/21/2017 12:42 PM
Formatted Table

Matthew Wozniak 8/9/2017 10:43 PM
Deleted: [a]

Matthew Wozniak 8/10/2017 4:36 PM
Deleted: -

Allison Steiner 8/21/2017 12:43 PM
Formatted: Font:Not Bold

families are organized into the four PFTs, with the two aggregated tree PFTs denoted as DBF and ENF. Regression slope (days/ºC) and coefficient of determination are provided for both sDOY and eDOY (slope/R^2).

Allison Steiner 8/21/2017 12:43 PM
Comment [2]: No longer shaded… did you mean to exclude? Either add back in or modify the caption.

Matthew Wozniak 8/23/2017 11:40 AM
Deleted: in gray shading

Allison Steiner 8/21/2017 12:43 PM
Formatted: Font:Not Bold

Land cover class	NE	SE	MT	CA	PNW
Acer	6.79E+04	2.88E+04	1.89E+03	1.97E+02	3.09E+03
Alnus	3.37E+00	1.23E-01	6.49E+01	1.71E+02	9.56E+03
Betula	2.99E+04	2.68E+03	2.78E+02	2.64E+00	4.82E+02
Fraxinus	3.96E+04	1.10E+04	3.14E+03	3.94E+01	2.76E+02
Morus	3.99E+03	2.25E+03	3.89E+01	0.00E+00	0.00E+00
Platanus	3.18E+03	3.38E+03	1.33E+01	1.44E+02	0.00E+00
Populus	5.48E+04	1.23E+03	4.37E+04	1.96E+02	1.55E+03
Quercus	1.30E+05	2.25E+05	2.51E+04	2.82E+04	1.40E+04
Ulmus	4.96E+04	2.81E+04	1.37E+03	0.00E+00	0.00E+00
BELD DBF	3.79E+05	3.03E+05	7.56E+04	2.90E+04	2.90E+04
CLM DBF	6.67E+05	4.03E+05	1.72E+05	7.93E+03	4.18E+04
Cupressaceae	1.85E+04	2.11E+04	7.84E+04	9.64E+03	2.35E+04
Pinaceae	8.34E+04	1.58E+05	1.79E+05	2.95E+04	1.10E+05
BELD ENF	1.02E+05	1.79E+05	2.58E+05	3.91E+04	1.34E+05
CLM ENF	1.44E+06	4.26E+05	4.66E+05	4.57E+04	5.34E+05

Table 2: Total spatial coverage (km^2) of tree taxa and PFTs from BELD and CLM land cover datasets in the 5 U.S. subregions (Northeast, NE; Southeast, SE; Mountain, MT; California; Pacific Northwest, PNW). All individual tree taxa are from the BELD database. BELD DBF and ENF land cover are the sums of the land cover of the taxa belonging to each PFT.

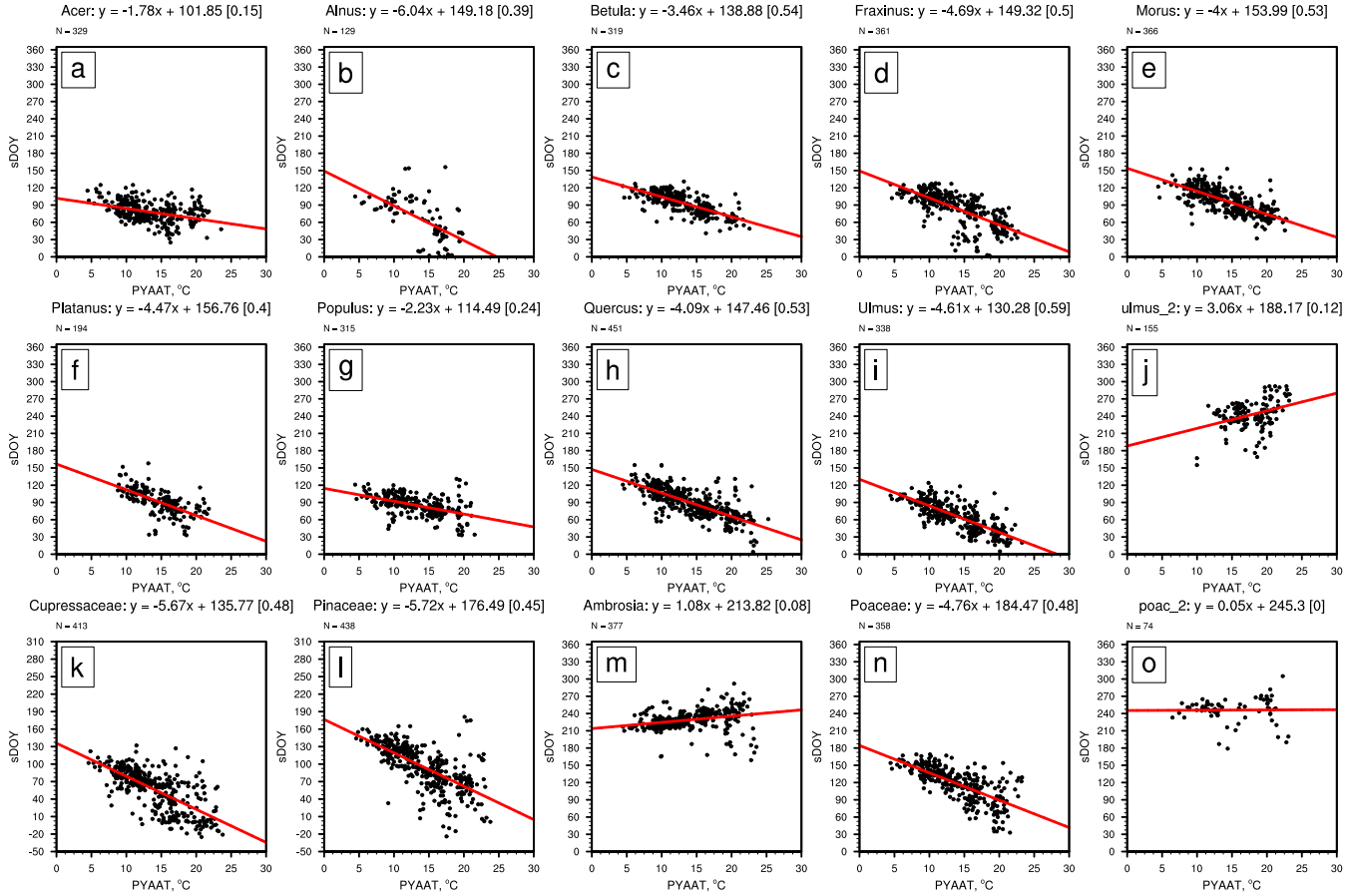


Figure S1. Linear regressions of the phenological relationship between start day-of-year (sDOY) and previous year annual average temperature (PYAAT) for all taxa. Each point signifies one station per year for pollen count data from 2003-2010 (total denoted as N). (j) ulmus_2 denotes a second (later) pollen peak in *Ulmus* and (o) poac_2 denotes a second (later) pollen peak in Poaceae.

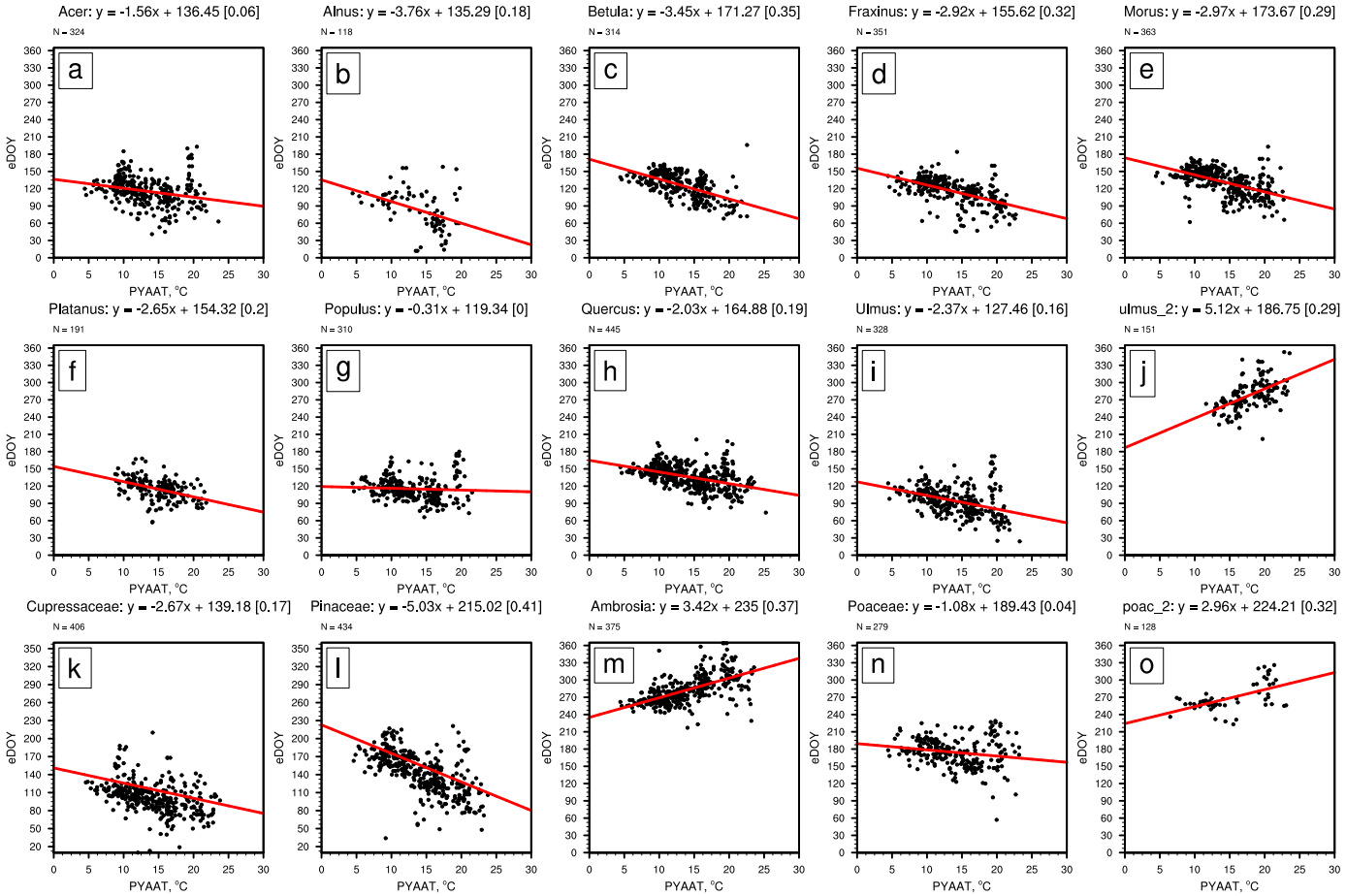


Figure S2. Same as Figure S1, but for end day-of-year (eDOY) versus PYAAT.

No.	City, State	Latitude	Longitude	Years	Acknowledgements
Northeast (38)					
1	Albany, NY	42.67	-73.80	2003-2010	David Shulan, M.D. Certified Allergy Consultants Albany, NY Certified Allergy Consultants
2	Armonk, NY	41.13	-73.71	2003-2010	Guy Robinson, PhD The Louis Calder Center Armonk, NY
3	Baltimore, MD	39.30	-76.61	2003-2010	Jonathon Matz, MD FAAAAI & David Golden, MD FAAAAI Dr. Golden and Dr. Matz, LLC Baltimore, MD
4	Boston, MA	42.35	-71.06	2010	Immunology Research Institute of New England Lawrence M. DuBuske, MD FAAAAI Boston, MA

5	Brooklyn, NY	40.65	-73.95	2004-2010	Clifford W. Bassett, MD & Mehdi Vesaghi, MD Long Island College Hospital Brooklyn, NY
6	Chelmsford, MA	42.60	-71.37	2003-2005	Julian Melamed, M.D. Chelmsford, MA
7	Cherry Hill, NJ	39.90	-75.00	2003-2010	Larchment Medical Center II Donald J. Dvorin, MD FAAAAI Cherry Hill, NJ
8	Chicago, IL	41.84	-87.68	2003-2010	John Shane, PhD McCrone Research Institute Chicago, IL
9	Dayton, OH	39.78	-84.20	2003-2010	Mr. Andy Roth RAPCA Dayton, OH
10*	Erie, Pennsylvania	42.13	-80.09	2003-2007, 2009-2010	Philip E. Gallagher, MD FAAAAI Allergy & Asthma Associates of Northeastern PA Erie, PA

11	Hamilton, Ontario	43.25	-79.1	2003-2005	Jason A. Ohayan, MD Hamilton, ON
12	Indianapolis, IN	39.78	-86.15	2003-2009	L.Y. Frank Wu, M.D. St. Vincent Professor Building Indianapolis, IN
13	Kansas City, MO	39.12	-94.55	2003-2007, 2009-2010	Jay Portnoy, MD FAAAAI Children's Mercy Hospital Kansas City, MO
14	La Crosse, WI	43.83	-91.23	2003-2010	N/A
15	Lexington, KY	38.043	-84.46	2003-2006,2008-2010	Beth Miller, M.D. University of Kentucky Asthma Allergy & Immunology Lexington, KY
16	Lincoln, NE	40.82	-96.69	2004-2009	Fred Keichel, MD Allergy, Asthma & Immunology Associates Lincoln, NE

17	London, ON	42.98	-81.25	2003-2005	James Anderson, MLT OSHTECH London, ON
18	Louisville, KY	38.22	-85.74	2003-2010	James L. Sublett, MD FAAAAI Family Allergy & Asthma Louisville, KY
19	Madison, WI	43.08	-89.39	2003-2010	Robert Bush, MD FAAAAI UW Medical School Madison, WI
20	Melrose Park, IL	41.90	-87.86	2003-2010	Joseph G. Leija, MD FAAAAI Dr. Joseph Leija Melrose Park, IL
21	Minneapolis, MN	44.96	-93.27	2010	Harold B. Kaiser, MD FAAAAI Clinical Research Institute Minneapolis, MN
22	Newark, NJ	40.72	-74.17	2003-2009	Alan Wolff, M.D. UMDNJ Newark, NJ

23*	New Castle, DE	39.62	-75.56	2005-2010	Michael McDowell Division of Air Quality, DNREC, State of Delaware New Castle, DE
24	New York, NY	40.08	-78.43	2003-2010	Guy Robinson, PhD Fordham College at Lincoln Center New York, NY
25	Niagrara Falls, ON	43.09	-79.09	2003-2005	Dr. Michael Alexander, MD Niagara Falls, ON
26	Olean, NY	42.08	-78.43	2003-2010	Fred Lewis, MD FAAAAI Olean, NY
27	Philadelphia, PA	40.01	-75.13	2003-2007, 2010	Donald J. Dvorin, MD FAAAAI Allergic Disease Associates, P.C. Philadelphia, PA
28	Pittsburgh, PA (2)	40.44	-79.98	2003-2010	David Skoner, MD FAAAAI Allegheny General Hospital Pittsburgh, PA

29	Rochester, NY	43.17	-77.62	2003-2010	Donald W. Pulver, MD FAAAAI Allergy, Asthma & Immunology of Rochester Rochester, NY
30	Salem, MA	42.53	-70.87	2007-2010	Paul Hannaway, M.D. Salem, MA
31	St. Claire Shores, MI	42.49	-82.89	2003-2005, 2009-2010	Andrew I. Dzul, MD Lakeshore Ear Nose & Throat Center St. Clair Shores, MI
32	St. Louis, MO	38.64	-90.24	2003,2006-2010	Mr. Wayne Wilhelm St. Louis County Health Department Berkeley, MO (St. Louis)
33	Washington, DC	38.91	-77.02	2003-2010	Susan E. Kosisky MA Walter Reed Army Medical Ctr. Washington, DC
34	Waturbury, CT	41.56	-73.04	2003-2010	Christopher Randolph, MD FAAAAI Waterbury, CT

35	Waukesha, WI	43.01	-88.24	2007-2009	Walter Brummund, MD, PhD, FAAAAI Allergy & Asthma Centers, S.C. Waukesha, WI
36	Wauwatosa, WI	43.06	-88.03	2003-2006	Walter Brummund, MD, PhD, FAAAAI Allergy & Asthma Centers, S.C. Waukesha, WI
37	York, PA	39.96	-76.73	2003-2010	Michael S. Nickels, MD PhD Allergy and Asthma Consultants, Inc. York, PA
Southeast (33)					
1	Atlanta, GA	33.76	-84.42	2003-2010	Santley M. Fineman, MD MBA FAAAAI Atanta Allergy and Asthma Marietta, GA (Atlanta)
2	Austin, TX	30.31	-97.95	2003-2005	Kim T. Hovanky, MD FAAAAI & Sheila M. Amar, MD FAAAAI, FACAAI Allergy & Asthma Center of Georgetown Austin, TX (Georgetown)

3	Baton Rouge, Louisiana	30.45	-91.13	2003-2005	James M. Kidd III Kidd Allergy Clinic Baton Rouge, LA
4	Birmingham, AL	33.53	-86.80	2010	Weilly Soong, MD FAAAAI Birmingham-Southern College/Alabama Allergy & Asthma Center Birmingham, AL
5	Charlotte, NC	35.20	-80.83	2003-2007	John T. Klimas, MD FAAAAI Carolina Asthma and Allergy Center Charlotte, NC
6	College Station, TX	30.60	-96.31	2003-2010	David R. Weldon, MD FAAAAI, FACAAI Scott & White Clinic College Station, TX
7	Corpus Christi, TX	27.69	-97.29	2003-2005	Gary L. Smith, M.D. Corpus Christi, TX
8	Dallas, TX	32.79	-96.77	2003-2010	Jeffrey Adelglass, M.D. Dr. Jeffrey Adelglass Dallas, TX

9	Fargo, ND	46.88	-96.82	2003-2010	Dan Dalan, MD FAAAAI Allergy & Asthma Care Center Fargo, ND
10	Flower Mound, TX	33.03	-97.09	2006-2010	Marie H Fitzgerald, MD North Texas Pollen Station Flower Mound, TX
11	Fort Smith, AK	35.37	-94.38	2003-2006	N/A
12	Georgetown, TX	30.65	-97.69	2006-2010	Kim T. Hovanky, MD FAAAAI & Sheila M. Amar, MD FAAAAI, FACAAI Allergy & Asthma Center of Georgetown Austin, TX (Georgetown)
13	Greenville, SC	34.84	-82.37	2003-2010	Neil L Kao MD FAAAAI Allergic Disease and Asthma Center Greenville, SC
14	Houston, TX	29.77	-95.39	2005-2010	Mr. Tony Huynh City of Houston Houston, TX

15	Huntsville, AL	34.71	-86.63	2003-2010	Ms. Debra Hopson Natural Resources & Environmental Management Huntsville, AL
16	Knoxville, TN	35.97	-83.95	2003-2010	Michael Miller, MD FAAAAI Allergy, Asthma and Immunology Knoxville, TN
17	Little Rock, AK	34.72	-92.35	2004-2010	Karl V Sitz, MD Little Rock Allergy & Asthma Clinic Little Rock, AR
18	Miami, FL	25.78	-80.21	2003-2005	Elene Ubals, MD & Richard Schiff, MD, PhD Miami, FL
19	New Orleans, LA	30.07	-89.93	2008-1010	W Edward Davis MD MS MBA MMM Ochsner Clinic Foundation New Orleans, LA
20	Ocala, FL	29.19	-82.13	2003-2005	Dr. Karl M Altenburger Allergy and Asthma Care of Florida Ocala, FL

21	Oklahoma City, OK (2)	35.47	-97.51	2003-2010	Warren V. Filley, MD FAAAAI OK Allergy Asthma Clinic, Inc. Oklahoma City, OK
22	Omaha, NE	41.26	-96.01	2003-2010	N/A
23	Orlando, FL	28.51	-81.37	2006-2007	Bruce A. Hornberger, MD FAAAAI Allergy & Asthma Center of East Orlando Orlando, FL
24	Oxford, AL	33.61	-85.84	2010	Robert Grubbe, MD Allergy & Asthma Center, LLC Oxford, AL
25	Rogers, AK	36.33	-94.13	2006-2010	Curtis L. Hedberg, MD FAAAAI Hedberg Allergy & Asthma Center Rogers, AR (Fort Smith)
26*	Sarasota, FL	27.34	-82.53	2003-2010	Mary Jelks, MD FAAAAI Sarasota, FL

27	Savannah, GA	32.02	-81.13	2008-2010	Brad H. Goodman, M.D. & Bruce D. Finkel, M.D. Coastal Allergy & Asthma, P.C. Savannah, GA
28	Tallahasseee, FL	30.46	-84.28	2003-2005	Ronald Saff, M.D. Tallahassee, FL
29	Tampa, FL	27.96	-82.48	2003-2010	Richard Lockey, MD FAAAAI University of South Florida Tampa, FL
30	Tulsa, OK	36.13	-95.92	2003-2010	Estelle Levetin, PhD FAAAAI University of Tulsa Tulsa, OK
31	Waco, TX (2)	31.57	-97.18	2003-2010	N.J. Amar, MD FAAAAI Allergy and Asthma Center Waco, TX
Mountain (8)					

1	Bismarck, ND	46.81	-100.77	2003-2005, 2010	Dan Dalan, M.D. North Dakota DOH East Lab Bismarck, ND
2	Colorado Springs (2)	38.86	-104.76	2003-2010	William Storms, MD FAAAAI The William Storms Allergy Clinic Colorado Springs, CO
3	Draper, UT	40.52	-111.86	2009-2010	Duane J. Harris, MD FAAAAI Intermountain Allergy & Asthma Clinic Draper, UT
4*	Missoula, MT	46.87	-114.01	2006-2007, 2009-2010	Emily Weiler U of Montana, Ctr for Environmental Health Sciences Missoula, MT
5	Salt Lake City, UT	40.78	-111.93	2003-2004, 2008	N/A
6	Scottsdale, AZ	33.69	-111.87	2006-2007	Michael E. Manning, M.D. Scottsdale, AZ

7*	Twin Falls, ID	42.56	-114.46	2003-2010	N/A
California (13)					
1	La Jolla, CA	32.84	-117.26	2003-2010	Robert Reid, Jr, MD Scripps Memorial Hospital LaJolla, CA
2	Orange, CA	33.81	-117.82	2003-2010	Sherwin A. Gillman, M.D. Children's Hospital of Orange County-Pediatric Subspecialty Faculty Orange, CA
3	Pasadena, CA	34.16	-118.14	2004-2005, 2010	Philip Taylor, PhD The Pollen Group Pasadena, CA
4	Pleasanton, CA	37.67	-121.89	2003-2010	N/A

5	Reno, NV	39.54	-119.82	2003-2006, 2009-2010	Leonard Shapiro, MD FAAAAI Allergy & Asthma Associates Sparks, NV
6	Roseville, CA	38.76	-121.29	2006-2010	Sunil P. Perera, MD FAAAAI Allergy Medical Group of the North Area Roseville, CA (Sacremento)
7	Sacramento, CA	38.57	-121.47	2003-2006	Sunil P. Perera, MD FAAAAI Allergy Medical Group of the North Area Roseville, CA (Sacremento)
8	Salinas, CA	36.68	-121.64	2003	Steven S. Prager, M.D. Salinas Allergy Clinic Salinas, CA
9	San Diego, CA	32.82	-117.14	2006, 2008-2010	Robert T. Reid, MD Erik and Ese Banck Clinical Research Center San Diego, CA
10	San Jose, CA (2)	37.30	-121.85	2003-2010	Theodore Chu, MD FAAAAI San Jose, CA

11	Santa Barbara, CA	34.43	-119.72	2003-2010	Myron Liebhaber, M.D. Sansum - Santa Barbara Medical Foundation Clinic Santa Barbara, CA
12	Stockton, CA	37.97	-121.31	2009-2010	Gregory W. Bensch, MD FAAAAI & George W Bensch, MD FAAAAI Allergy, Immunology and Asthma Medical Group Stockton, CA
Pacific NW (4)					
1	Crescent City, CA	41.76	-124.20	2009-2010	N/A
2*	Eugene, OR	44.05	-123.11	2003-2010	Kraig W. Jacobson, MD, FAAAAI Allergy & Asthma Research Group Eugene, OR
3	Seattle, WA	47.62	-122.35	2003-2010	N/A

4*	Vancouver, WA	45.63	-122.64	2003-2010	Raymond Brady, M.D. and Joseph Hassett, M.D. Vancouver, WA

Table S1. AAAAI pollen counting stations included in development and evaluation of model. Categorized by U.S. subregions Northeast, Southeast, Mountain, California, and Pacific Northwest. Sites equipped with two samplers are denoted in parentheses (ex. Pittsburgh, PA (2)). Station numbers with asterisk (*) are collocated with DayMet temperature for phenological regressions.

1062

Taxon	Maximum P_{avgmax} (grains m^{-3})	Average P_{avgmax} (grains m^{-3})
Acer	535	212
Alnus	176	73
Ambrosia	379	193
Arecaceae	27	13
Artemisia	76	36
Asteraceae	64	28
Betula	564	232
Carpinus/Ostrya	71	29
Carya	126	54
Celtis	155	65
Chenopodiaceae/Amaranthaceae	80	29
Corylus	13	5
Cupressaceae	1949	960
Cyperaceae	37	12
Eupatorium	27	13
Fagus	66	17
Fraxinus	462	227
Graminae / Poaceae	197	98
Juglans	71	34
Ligustrum	13	6
Liquidambar	109	44
Morus	940	398
Myrica	50	25
Olea	197	43
Other Grass Pollen	85	42
Other Tree Pollen	863	332
Other Weed Pollen	160	51
Pinaceae	645	328
Plantago	33	13
Platanus	191	103
Populus	367	158
Prosopis	34	16
Pseudotsuga	7	3

Quercus	2106	1081
Rumex	59	23
Salix	123	58
Tilia	19	6
Tsuga	12	5
Typha	19	10
Ulmus	556	249
Unidentified Pollen	198	57
Urticaceae	86	36

Table S2. ***Maximum and average*** P_{avgmax} used for selecting pollen taxa for this study. Shaded taxa are selected taxa based on minimum of 1) 100 grains m^{-3} for the maximum P_{avgmax} and 2) 70 grains m^{-3} for the average P_{avgmax}.

Matthew Wozniak 8/9/2017 11:27 AM
Deleted: SI