Allergenic pollen production across a large city for common ragweed (*Ambrosia artemisiifolia*)

Daniel S.W. Katz*, Stuart A. Batterman
School of Public Health, University of Michigan – Ann Arbor, Ann Arbor, MI, USA

Abstract

Predictions of airborne allergenic pollen concentrations at fine spatial scales require information on source plant location and pollen production. Such data are lacking at the urban scale, largely because manually mapping allergenic pollen producing plants across large areas is infeasible. However, modest-sized field surveys paired with allometric equations, remote sensing, and habitat distribution models can predict where these plants occur and how much pollen they produce. In this study, common ragweed (*Ambrosia artemisiifolia*) was mapped in a field survey in Detroit, MI, USA. The relationship between ragweed presence and habitat-related variables derived from aerial imagery, LiDAR, and municipal data were used to create a habitat distribution model, which was then used to predict ragweed presence across the study area (392 km²). The relationship between inflorescence length and pollen production was used to predict pollen production in the city. Ragweed occurs in 1.7% of Detroit and total pollen production is $312 \times 10^{12}$ pollen grains annually, but ragweed presence was highly heterogeneous across the city. Ragweed was predominantly found in vacant lots (75%) and near demolished structures (48%), and had varying associations with land cover types (e.g., sparse vegetation, trees, pavement) detected by remote sensing. These findings also suggest several management strategies that could help reduce levels of allergenic pollen, including appropriate post-demolition management practices. Spatially-resolved predictions for pollen production will allow mechanistic modeling of airborne allergenic pollen and improved exposure estimates for use in epidemiological and other applications.

Keywords

Aerial imagery; Allergic rhinitis; Habitat distribution modeling; LiDAR; Urban ecology; Vacant lots

1. Introduction

Allergenic pollen concentrations can vary by orders of magnitude across spatial scales of hundreds to thousands of meters (Bricchi, Frenguelli, & Mincigrucci, 2000; Katz & Carey, 2014; Weinberger et al., 2018; Werchan et al., 2017), but this spatial heterogeneity is rarely...
accounted for in epidemiological or other studies of allergenic pollen (Lovasi et al., 2013). Pollen concentrations are not estimated at finer scales primarily because the location and abundance of pollen producing plants are generally unknown (Zink et al., 2016). While field surveys of plants in small areas can help to explain local airborne pollen concentrations (Bricchi et al., 2000; Charalamopoulos, Lazarina, Tsiripidis, & Vokou, 2018; Nowak, Szymanśka, & Grewling, 2012), it is not logistically feasible to apply this approach to large areas. Airborne pollen concentrations have been predicted using land use cover and land use regression (Hjort et al., 2015; Maya-Manzano et al., 2017), but land cover classifications are coarse proxies for habitat suitability and may be ill-suited for use in mechanistic models that require estimates of pollen production. Fine grain models of plant presence and pollen production could provide emission or source information for atmospheric dispersion models, which could dramatically improve predictions of airborne pollen concentrations (Skjøth et al., 2013).

One of the most important allergenic pollen producing plants in North America, Europe, and Asia is common ragweed (Ambrosia artemisiifolia L.), an early successional annual plant that is expected to become even more problematic due to climate change (Lake et al., 2016). Ragweed is a prolific producer of allergenic pollen, and approximately 15–20% of Americans are allergic to it (Trends, 2011; Salo et al., 2011). Ragweed pollen exposure results in a variety of respiratory symptoms (Caillaud et al., 2014) and is associated with emergency room visits for asthma (Breton, Garneau, Fortier, Guay, & Louis, 2006; Makra, Matyasovszky, Bálint, & Csépe, 2014; Zhong et al., 2006) and allergy medication sales (Motreff et al., 2014).

Most efforts to predict ragweed plant presence and abundance have occurred at national (Dullinger, Kleinbauer, Peterseil, Smolik, & Essl, 2009) or continental scales (Chapman, Haynes, Beal, Essl, & Bullock, 2014; Cunze, Leiblein, & Tackenberg, 2013; Leiblein-Wild, Steinkamp, Hickler, & Tackenberg, 2016), or have been extrapolated from airborne concentrations of ragweed pollen (Karrer et al., 2015; Thibaudon, Šikoparija, Oliver, Smith, & Skjøth, 2014). These efforts are important for understanding large scale processes relevant to long distance transportation. In addition, strong evidence exists that ragweed pollen concentrations vary spatially within cities and that variation has been tied to local plant abundance (Katz & Carey, 2014). While fine scale predictions of ragweed plant presence have been explored in agricultural settings using remote sensing data (Auda et al., 2008; Csornai et al., 2009), the feasibility of estimating plant location and pollen production in complex urban areas is unknown.

Habitat distribution models (and related models such as species distribution models and occupancy models) can provide the basis for predicting the presence and abundance of plants within urban areas (e.g., Chance et al., 2016). Remote sensing data (e.g., LiDAR, aerial imagery, and satellite imagery) is becoming increasingly available and is now used extensively to inform species distribution models (He et al., 2015). When paired with field surveys, remote sensing can successfully predict plant presence at scales of tens to hundreds of meters (Akasaka, Osawa, & Ikekami, 2015; Tuanmu et al., 2010; Wang, Zachmann, Sesnie, Olsson, & Dickson, 2014). Remote sensing also can inform predictions in inaccessible areas, such as private property in urban areas. For species that inhabit
developed areas, socioeconomic data can also be included in habitat suitability models; this is commonly applied for disease vectors like mosquitos (Sallam, Fizer, Pilant, & Whung, 2017), but rarely for plants. While the techniques and remote sensing data are available, we are unaware of prior studies predicting allergenic pollen producing plant locations and pollen production in urban areas.

The objective of this study is to demonstrate, apply, and evaluate a method for predicting plant presence and pollen production across a large urban area. Ragweed plant locations and flower abundance were recorded in a field study, and then these data were combined with remote sensing and municipal data to create a habitat distribution model for ragweed presence. Pollen production is estimated at high spatial resolution using flower abundance and allometric equations. Predictions across the study area are evaluated using cross validation and simulation.

2. Methods

2.1. Field surveys

We conducted field surveys across three years to determine locations and pollen production of ragweed plants in Detroit, Michigan, USA. Surveys were conducted in late summers (August and September) of 2016, 2017, and 2018, covering an area of 2.58, 1.14, and 1.42 km², respectively (Fig. 1A). In 2016, sampling neighborhoods were determined by dividing the city into equal sections, and a randomly selected section was surveyed each field day using belt transects. In 2017 and 2018, a portion (7%) of areas surveyed in 2016 were resurveyed, and additional neighborhoods in the vicinity of airborne pollen sampling sites from a related experiment were also surveyed. Transects were walked in a predetermined manner through unobstructed areas (e.g., sidewalks, vacant lots, and parks; Fig. 1B). The purpose of these transects was to determine ragweed distributions within each potential habitat type, and not to survey a representative amount of each type of ragweed habitat.

Ragweed plants taller than 20 cm and within 10 m of the transect center line were measured. Variables recorded in 2016 included: plant height (defined as the distance between the ground and the tallest apical meristem or flower tip, measured to a precision of 1 cm using a meter stick); stem diameter (one cm above ground, measured with a digital caliper to 0.01 cm); and the cumulative length of male inflorescence, recorded to the nearest 10 cm; for plants with greater than 200 cm of flowers, this was extrapolated based on measurements from a representative fraction of the plant (one quarter or one eighth, depending on plant size). Stem diameter was not recorded in 2017; in 2018, only plant coordinates were recorded. A GPS receiver (GNSS Surveyor, Bad Elf, Tariffville, CT, USA) collected location measurements every second during sampling. Coordinates for individual plants were recorded using the GIS Collector application (ESRI, Redlands, CA, USA) on an Ipad Mini 2 (Apple, Cupertino, CA, USA). The general location was determined using coordinates from the GPS, and specific plant locations were recorded using the ESRI basemap imagery to an accuracy of ~ 1 m. When six or more ragweed plants were found in a homogenous area, the area was delineated with a polygon, the same measurements were taken on six randomly selected plants within the polygon, and the number of ragweed plants within the polygon was counted.
Survey data were summarized at the spatial resolution of 4 m (4 m × 4 m pixels). This resolution was selected as smaller pixels would increase both the potential for spatial misalignment and the computational demand, whereas larger pixel sizes would have obscured locally relevant differences in land use and created more heterogeneous (“mixed”) pixels. Only pixels entirely within the surveyed area (< 10 m from the transect centerline) were included.

### 2.1. Remote sensing

Remote sensing data were used to predict ragweed presence (Fig. 2). Similar spectral clusters were identified using publicly available aerial imagery collected in summer 2016 (National Agricultural Imaging Program, NAIP 2016; 4 bands, 0.5 m) and k-means unsupervised classification (50 classes) with the RSToolbox package (Leutner & Horning, 2018). Classification categories (classes) were associated with particular landscape features (e.g., dark pavement or dense vegetation) and are described in SI 1. The proportion of each 16 m² pixel occupied by each class was used.

Land cover was classified as trees, buildings, or ground using LiDAR (1.0 ppm QL2 provided by the Michigan Statewide Imagery Program, collected spring 2017) and aerial imagery (described above). A digital elevation model (DEM: ground), a surface digital model (SDM; first returns), and a digital terrain model (DTM; last returns) were created for the study area using the R package lidR (Roussel & Auty, 2017) and a pit-free algorithm (Khosravipour, Skidmore, Isenburg, Wang, & Hussin, 2014) implemented in lidR with processing through a 9-cell median filter. The difference between the SDM and DTM (i.e., the distance that points penetrate into an object) provide a good operational basis for distinguishing between trees and buildings (O’Neil-Dunne, MacFaden, & Royar, 2014). This classification was supplemented with NDVI values derived from the NAIP imagery to distinguish between trees and buildings. Pixels with a height below 3.05 m (10’) were classified as ground (including shrubs, saplings, and sheds). The most common cover type within the study area was ground (64%), followed by trees (21%) and buildings (15%). Classification accuracy was assessed using 200 randomly generated control points, which were manually classified using high resolution aerial imagery and LiDAR. The overall accuracy of the cover classification rate was 98% (95% CI: 0.95–0.99; SI 2). The most common land cover classification within each 16 m² pixel was used.

Demolished structures were classified using building footprint shape files created from aerial imagery collected in 2010 (provided by the Southeast Michigan Council of Governments). Buildings were assumed to have been demolished when the height of most of the area in the building footprint in 2017 was < 2 m in the 2017 LiDAR dataset. This is a conservative approach as it excludes buildings that were mostly overhung by trees. Because the demolition process can affect areas beyond the building’s footprint, the demolished buildings were buffered at distances of 8, 16, 32 and 64 m. The Detroit Demolition Program demolished 13,290 structures in Detroit by the end of 2017. While publicly available, this database does not include the small cities of Highland Park and Hamtramck (located within the city of Detroit), privately funded demolitions, or the exact location of demolished buildings, so this database is not used here. Demolition classification accuracy was assessed
by manually extracting the demolition status of 100 stratified random control points from aerial images taken in 2010 through the MiSAIL program and from high resolution aerial images taken in spring and fall 2017 provided by NearMap (Sydney, Australia). The analysis of demolished structures between 2010 and spring 2017 had an overall accuracy of 99% SI 3. 9% of the study area was within 32 m of a demolished structure, and 16% of the study area was occupied by vacant lots.

2.2. Municipal data

Ragweed occurrence and abundance are determined in large part by human activities. Socioeconomic data were included at two scales: census blocks from the 2016 US census; and property parcel records made available by the city of Detroit. Census block boundaries and the 2016 American Community Survey (Census, 2016) were downloaded and converted to rasters to align with our other datasets. Potentially relevant variables on income (household income, per capita income), demography (total population), and building status were obtained (vacant housing units, median rent, median house value, median year built).

Parcel data included shape files with information on parcel location, size, assessed land value, and improved value (i.e., the assessed value of built structures on the parcel). Parcels with a zero improved value were considered to be vacant. For the cities of Highland Park and Hamtramck, where these data were not publicly available, we selected all parcels where less than 10% of land cover was classified as a building, manually reviewed the aerial imagery of the parcels, and flagged vacant lots. Parcel boundaries often did not extend to the edge of sidewalks, so boundaries of all parcels were extended by 4 m (overlapping between parcels was not permitted).

Additional land use types included parks and roads. Park boundaries were downloaded from the City of Detroit open data portal, and cemeteries and golf courses in Detroit were digitized and manually added to this classification. Road centerlines were downloaded from the same portal; these were then buffered by 4 m on either side. This is a conservative labeling of roads because many roads are wider than 8 m. (Wider buffers would have exceeded the true road edge on many residential streets.) Roads and parks were coded as binary variables.

2.3. Statistical analyses

Moran’s I and semivariograms showed substantial spatial auto-correlation for both dependent and independent variables. To better satisfy assumptions of independence between pixels, the dataset was thinned prior to analysis, retaining only one of 16 pixels (6.25%; minimum distance between pixels was 16 m). During fieldwork, we observed that parcels were the primary unit of management and land use history. Ragweed presence within a parcel remained auto-correlated, as were exploratory model residuals. To further reduce spatial autocorrelation, only a single pixel from each parcel was included in the analysis; semivariograms and Moran’s I of final model residuals are reported in SI 4. The final sample size for the model was 7908.
Logistic regression was used to analyze how land cover, land use, image characteristics, and socioeconomic data affected ragweed presence. Parameter estimation used a Bayesian approach, which allowed incorporation of missing data, e.g., land values from Hamtramck and Highland Park (Gelman & Hill, 2007). Ragweed presence for each pixel $i$ was estimated as a function of covariates $X_i$, where $X_i$ is the vector of covariates associated with each pixel. After exploring several model structures (e.g., conditional autoregressive models, spatially nested models), we ultimately chose a simple logistic regression method. Variable selection was conducted in step-wise fashion with Deviance Information Criterion (Plummer, 2003) until an additional variable did not reduce DIC by more than 3. (Tested models and DIC are provided in SI 5.) The final model included an intercept for each general land cover type ($\beta$: ground, trees, and buildings), fixed effects for several binary land use variables ($\delta$: road, demolition within 16 m, demolition within 64 m, and vacant lot), and several image classifications categories ($\alpha$: light pavement, dark pavement, light vegetation, sparse vegetation), but did not include any of the census variables, or the other 46 image classification categories:

$$\text{Observed presence } i \sim \text{Bernoulli (presence}_i)$$

$$\text{Logit(presence}_i) = \alpha_1 \times \text{dark pavement}_i + \alpha_2 \times \text{mixed vegetation}_i + \alpha_3 \times \text{light pavement}_i + \alpha_4 \times \text{sparse vegetation}_i + \delta \text{ vacant}_i + \delta \text{ demolition16}_i + \delta \text{ road}_i + \delta \text{ demolition64}_i + \beta \text{ cover}_i$$

Fixed effect coefficients ($\alpha^*$, $\delta^*$) and intercepts ($\beta^*$) were drawn from non-informative prior distributions: $\alpha^*$, $\delta^*$, and $\beta^*$: Normal (0, 1000). (Code is available in SI 6.)

Model coefficients were considered statistically significant if their 95% CI did not include zero. Posterior densities of the parameters were obtained by Gibbs sampling using JAGS 4.2 (Plummer, 2003) via the rjags package in R (Plummer 2014) and the runjags package in R (Denwood, 2016). Convergence occurred after 500–1000 iterations; chains were inspected visually and assessed with the potential scale reduction factor statistic. Posterior parameter values were based on post-convergence results. Occupancy models were assessed using AUROC curves and precision recall curves (Fawcett, 2006; Grau, Grosse, & Keilwagen, 2015) based on test data (the surveyed areas were divided in to 128 m × 128 m areas; 20% of these areas were used for test data). Predictions were created by calculating all combinations of independent variables and simulating ragweed presence for each scenario. The mean and standard deviation of predicted ragweed presence for each scenario was then applied across a raster stack containing all predictor variables. Open water (i.e., the Detroit River) was assumed to have no ragweed present.

### 2.4. Pollen production

Pollen production per plant was estimated based on the cumulative inflorescence length using locally-developed allometric equations reported elsewhere (Bankowski & Katz, 2018). Unlike other published estimates of ragweed pollen production as a function of inflorescence length, this model was derived from wild grown plants exposed to representative growing conditions (including competition, herbivory, and mowing). The following model was fitted to the total total pollen production of 33 individuals (by counting the number of flowers per inflorescence, anthers per flower, and pollen grains per anther):
\[ p = 624,018x - 6,025,129 \]

where \( p \) is total pollen production, and \( x \) is the total inflorescence length (restricted to values > 10, the minimum inflorescence length recorded in our study; \( R^2 = 0.91, p < 0.0001 \)). This equation provides similar predictions as an equation based on ragweed plants grown in a greenhouse and subjected to a mowing treatment (Simard & Benoit, 2011).

Although several variables were correlated with ragweed flowers per pixel (where ragweed was present), these associations were weak and statistically insignificant and models yielded little explanatory capacity (\( R^2 < 0.10 \)). Total pollen production was estimated by multiplying the probability of ragweed occurrence in each pixel by the average amount of pollen produced per pixel. To compare differences in pollen production at some of the spatial scales potentially relevant to urban scale applications and dispersion modeling of pollen, the average pollen production per m\(^2\) was calculated and mapped at spatial resolutions of 0 (pixel level), 100, 250, 500, 1000, and 2000 m.

### 3. Results

#### 3.1. Field survey

Ragweed was present in 2.4, 2.8, and 0.7% of the 16 m\(^2\) pixels surveyed in 2016, 2017, and 2018, respectively. In 2016 and 2017, a total of 24,731 and 4486 plants, respectively, were mapped; in 2018, over 817 plants were mapped (the total number of plants in patches were not counted in 2018). In 2016, measurements taken on 2065 individual plants showed a mean height of 39.6 cm (range: 20–191, SD: 19.6), mean diameter of 4.7 mm (range: 0.5–37.0, SD: 3.4), and inflorescence length of 75.7 cm (range: 0–6400, SD: 298.8). (The 2016 campaign also noted 1731 giant ragweed plants (\( A. \) \textit{trilida} L.) that were excluded from the analyses but no other species of ragweed were observed.) In 2017, measurements were taken on 764 individual plants; these had a mean height of 37.1 cm (range: 20–146, SD: 15.5), and mean inflorescence length of 98 cm (range: 0–3900, SD: 248.9). Measurements on individual plants were not collected in 2018. Across the study period, 75% of the ragweed was found in vacant lots, 48% in pixels that had a demolition within 32 m, and 41% in pixels that were both vacant and near a demolished structure. Ragweed occurred in 3.3% of parcels surveyed. Of the parcels with ragweed, 75.9% were vacant, 64.5% contained a demolished structure, and 50.4% were vacant lots where a demolition had occurred (SI 7). These results may not be representative of other cities given the magnitude of Detroit’s demolition program, however, they indicate the importance of addressing disturbances to minimize ragweed prevalence (discussed later).

To examine potential differences in ragweed presence and flower abundance over the study period, a subsample of 7466 pixels surveyed in all years were compared. This subsample had similar ragweed presence in 2016 (average of 7.8%) and 2017 (6.0%), but substantially less ragweed in 2018 (1.5%). This is similar to differences between modeled ragweed occurrence rates in these years (presented below). This variation may result in part from higher mowing rates and maintenance in 2018 (Katz, personal observation). There were also more demolitions in 2015 (4014) and 2016 (3201) than in 2017 (2508), and there may also have been differences in demolition practices, weather, and other factors. Despite the annual
variation in total ragweed occurrence, certain neighborhood appear to have consistently more ragweed (Katz & Carey, 2014).

3.2. Habitat model results

Ragweed was predicted to occur across 1.7% of the city in 2016 and 2017, and probabilities for individual pixels ranged from 0 to 38%. The probability varied substantially by land cover, land use, and image classification (Table 1, Fig. 3, Fig. 4). The predicted probability of ragweed (pixel-level) was $0.78 \pm 0.24\%$ (SD) in bare ground pixels, $2.96 \pm 0.73\%$ in vacant lots, $1.61 \pm 0.43\%$ within 64 m of a demolition, and $4.18 \pm 1.21\%$ within 16 m of a demolition. Probabilities were low for pixels classified as roads ($0.38 \pm 0.21\%$), buildings ($0.0011 \pm 0.009\%$), or trees ($0.23 \pm 0.10\%$). The presence of ragweed in buildings is an artifact due to small amounts of ground being classified as buildings (cover type was assigned by pixel majority) but ragweed did grow through pavement cracks and in accumulated organic matter in road margins. Ragweed presence was higher in areas where the aerial imagery was classified as thin vegetation (mixed vegetation and soil pixels as seen in Fig. 1C, $12.10 \pm 9.87\%$), and was negatively associated with the dark (greater than $0.01 \pm 0.01\%$) and light pavement ($0.01 \pm 0.02\%$) image categories. Combinations of these variables influenced ragweed probability, for example, ragweed was predicted to occur in $14.3 \pm 2.5\%$ of pixels that were both within 16 m of a demolition and vacant. The model included a fixed effect for which was predicted to have far lower ragweed occurrence across the city ($0.4\%$).

The AUROC of the model test data was 0.85 (SI 8), which can be interpreted as the model ranking a pixel where ragweed was present higher than a pixel where ragweed was absent 85% of the time in the test data. Precision-recall curves are useful diagnostics for models of rare events; the AUC of the precision recall curves was 0.19 (SI 8). Thus, this model is suited for understanding general patterns of ragweed presence, even though predictions for a specific pixel may not be accurate.

3.3. Ragweed pollen production

In pixels where ragweed was observed, average cumulative measured ragweed inflorescence was 1310 cm, or 82 cm/m$^2$ (Supporting Information 10). Based on mean pollen production in pixels where ragweed was present (51,000,000 grains/m$^2$) and the average ragweed likelihood (1.7%), a total of $312 \times 10^{12}$ pollen grains were produced annually over the study area. This is equivalent to 1.36 cm/m$^2$ of ragweed inflorescence, 847,000 grains/m$^2$ per year across the City of Detroit. However, pollen production was very heterogeneous, and spatial differences in production rates are seen, even at large averaging scales (Fig. 5).

4. Discussion

The presence of an important allergenic pollen producing plant was predicted across a large urban area based on a field census, publicly available remote sensing data, and a habitat distribution model. Understanding plant locations and pollen productions rates is important: ecologically-based estimates of plant location and pollen production can inform the development of management strategies; and this information can provides the source
or emission data that allows the use of physically-based dispersion models to estimate airborne concentrations of allergenic pollen. The fusion of field censuses, remote sensing, habitat modeling, and allometric pollen production equations used here could be adapted to estimate pollen production for other areas or species. The method provides an alternative to the use of land use regression to predict pollen concentrations (e.g., Hjort et al., 2015) with potential advantages that its more mechanistic basis and higher spatial resolution may improve predictions and allow more accurate forecasts of the effectiveness of management strategies.

4.1. Ragweed presence

Ragweed presence varied considerably and predictably by land use, land use history, cover type, and the presence of pavement and roads. The analysis captured ragweed’s general habitat preferences (Smith, Cecchi, Skjøth, Karrer, & Šikoparija, 2013) and generally matched previous observations of urban ragweed populations (Katz, Connor Barrie, & Carey, 2014; Urbanowicz, Hutyra, & Stinson, 2018; Vincent & Bergeron, 1985). In particular, the link between building demolition and ragweed presence is well supported; ragweed germination is highest when seeds are at the soil surface (Pickett & Baskin, 1973). The presence of ragweed in areas near a demolition but that were not in a vacant lot might result from locally increased seed production and dispersal, or perhaps from disturbances that cross property lines.

4.2. Pollen production

When ragweed was present, models based on flower abundance could explain only a small fraction ($R^2 < 0.10$) of the variation in flower abundance. The available variables (e.g., land cover type, vacant lot status, land value, income, and location) also explained little of the variation; field observations suggest this is due to a lack of data on mowing, management type, seed bank composition, and inter-specific plant competition. Mowing can dramatically reduce ragweed pollen production (Simard & Benoit, 2011), and we occasionally found areas where large ragweed patches had recently been mowed, resulting in low pollen production even though many ragweed plants were present. Conversely, a total lack of mowing in vacant lots results in ragweed being outcompeted within a few years of soil disturbance (Katz et al., 2014). Other underlying variables that might affect ragweed abundance, e.g., seed bank, are also difficult to obtain at the municipal level.

Pollen production estimates (Fig. 5, panel 1) use the average pollen production per pixel and do not incorporate the substantial variation in that distribution (SI 9), resulting in considerable uncertainty at the pixel level. Total pollen production averaged across larger spatial scales (e.g., Fig. 5, panels 2–6) will be less affected by the variability of individual pixels. Fig. 5 shows the effect of variability at the pixel level on total pollen production for various distances from a specific location (i.e., a receptor) based on simulation methods reported in SI 10. If individual pixels are assumed to be independent, unexplained variation in pollen production at the pixel level has only a small effect on estimates of total pollen production at larger spatial scales, e.g., by 1000 m the coefficient of variation (COV) between simulations is 0.04; assuming pixels have the same pollen production, COV = 0.27. These bounding assumptions and results in SI 10 suggest that the reliability of the pollen
production surface is highest when averaged across distances of ≥1000 m but that results will be useful at distances down to 250 m. These spatial scales are similar to the spatial scales of ragweed pollen dispersion, most of which occurs at tens to hundreds of meters (Martin, Chamecki, Brush, Meneveau, & Parlange, 2009; Raynor et al., 1970; Raynor & Ogden, 1965; Raynor, Smith, Singer, & Ogden, 1961).

4.3. Field surveys in urban areas

Traditional approaches to field surveys can be difficult to implement in urban areas because substantial portions of cities are legally or logistically inaccessible. Randomly selecting sampling sites within the city, with the intention of obtaining a representative sample, would require permission from land owners. Instead, we establish relationships between spatial variables and ragweed presence, and use these relationships to predict pollen production across the city, including inaccessible locations. This approach assumes that the relationships between ragweed and explanatory variables (e.g., tree cover) are consistent between areas visited and other areas. This makes the use of imagery especially important as it can capture existing conditions (e.g., sparse vegetation, pavement, lawns), which might systematically vary between accessible and inaccessible areas.

4.4. Management recommendations and demolitions

Building demolitions (and the resulting vacant lots) are the best predictors of ragweed presence in Detroit. There are several post-demolition strategies that could reduce allergenic pollen exposure (Bullock, Chapman, Schafer, Roy, Haynes, Beal, & Tinch, 2012; Essl et al., 2015): the use of an ecologically diverse seed mixture containing plants that out-compete ragweed in the first year or two after demolition (Christopher Swan, personal communication); mowing lots several times a year (Simard & Benoit, 2011); and not mowing lots at all (not currently allowed by ordinance) (Katz et al., 2014). The health benefits of these and other control strategies could be evaluated using an atmospheric pollen dispersion model to quantify contributions of both nearby and distant ragweed plants coupled to a health impact or risk model (based on epidemiological data) to quantify the avoided adverse health impacts, such as asthma attacks and lost work/school days due to asthma aggravation.

Managing land to reduce ragweed pollen production may not be feasible across large areas. However, appropriate strategies applied to ragweed “hotspots” can be effective. For example, our results indicate that managing the top 1% of area in Detroit could reduce city-wide ragweed pollen production by 33%. Because identifying the worst parcels is labor intensive, ways to narrow the search are useful. For cities like Detroit where ragweed frequently occurs in vacant lots, efforts could focus on parcels that are both vacant and that had a recent demolition. In Detroit, this represents only 18% of all parcels, of which 52% are expected to contain ragweed.

4.5. Exposure estimates and health investigations

Previous epidemiological investigations of ragweed pollen (Breton et al., 2006; Caillaud et al., 2014; Darrow et al., 2012; Gleason, Bielory, & Fagliano, 2014; Héguy et al., 2008; Makra et al., 2014; Motreff et al., 2014) have used exposure estimates based on
airborne pollen measurements collected at a single point within a city. However, our results demonstrate substantial intraurban variation in ragweed pollen production, and prior work shows that ragweed abundance is reflected in airborne pollen concentrations (Katz & Carey, 2014). This suggests that these previous epidemiological investigations of ragweed may have been affected by exposure measurement errors. Indeed, authors of one study noted that poorer neighborhoods seemed especially affected (Breton et al., 2006). Intraurban variation in plant locations and pollen release are also relevant to other taxa including trees (Katz, Dzul, Kendel, & Batterman, 2019; Weinberger, Kinney, & Lovasi, 2015) and grasses (Hjort et al., 2015). Recently, Katz et al. (2019) showed that the use of a single measurement location to reflect pollen levels may not adequately reflect the timing of the pollen season, causing exposure measurement errors important to longitudinal studies.

Better methods to estimate airborne pollen concentrations are required for epidemiological analyses, to improve the pollen forecasting and alert systems, and to evaluate management options for reducing airborne pollen levels. The present study suggests that models of pollen production in combination with dispersion models can be used to estimate exposure and account for both spatial and temporal variability within an urban area. Such models may reduce exposure measurement error in epidemiological analyses since intra-urban variation in plant abundance is incorporated, and since the pixel-level uncertainty in plant location and pollen production will, to a large extent, be smoothed out by atmospheric pollen dispersion occurring over tens to hundreds of meters (Raynor et al., 1970). While the development and evaluation of such coupled models is beyond the present scope, this approach appears a promising research direction.

5. Conclusions

Mechanistically-based models of airborne pollen concentration require data on pollen production across large and often complex urban landscapes. This study modeled plant presence as a function of commonly available remote sensing and socio-economic datasets. The resulting fine-grained plant occurrence predictions were combined with allometric equations to estimate pollen production across the study area. The habitat model also provides insight to the focal species ecology and the development of appropriate strategies to target and remove this species. This approach could be implemented in other areas, as well as for other species of allergic importance.

6. Data availability

Data associated with this paper have been deposited in Dryad: https://doi.org/10.5061/dryad.bf08b1t.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.
Acknowledgements

This work was supported by the National Institute of Environmental Health through a NRSA postdoctoral fellowship (Grant Number F32 ES026477). It was also supported by the Michigan Institute for Clinical Health Research through the Postdoctoral Translational Scholars Program (Grant Number UL1 TR002240). S. Batterman also acknowledges support from grant P30ES017885 from the National Institute of Environmental Health Sciences, National Institutes of Health. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. We thank the students and research assistants that contributed to this work: John Kost, Victoria Bankowski, Floyd Watkins, and Yundi Yang. Additional thanks to Ines Ibanez, the Global Change Ecology lab, and Deborah Goldberg for advice, feedback and support.

References

Akasaka M, Osawa T, & Ikegami M (2015). The role of roads and urban area in occurrence of an ornamental invasive weed: A case of *Rudbeckia laciniata* L. Urban Ecosystems, 1021–1030. 10.1007/s11252-015-0466-4.

Auda Y, Déchamp C, Dedieu G, Blasco F, Dusit D, & Pontier EJ-L (2008). Détection des plantes envahissantes par télédétection: Un cas d’étude, l’ambroisie en région Rhône-Alpes, France. International Journal of Remote Sensing, 29(4), 1109–1124. 10.1080/01431160701355231.

Bankowski V & Katz D (2018). Estimates of common ragweed pollen production for urban ragweed plants. Retrieved from The University of Michigan Undergraduate Research Journal, 12, 27–32. http://www.umurj.org/.

Breton M-C, Garneau M, Fortier I, Guay F, & Louis J (2006). Relationship between climate, pollen concentrations of *Ambrosia* and medical consultations for allergic rhinitis in Montreal, 1994-2002. Science of the Total Environment, 370(1), 39–50. 10.1016/j.scitotenv.2006.05.022.

Bricchi E, Frenguelli G, & Mincigrucci G (2000). Experimental results about *Platanus* pollen deposition. Retrieved from Aerobiologia, 16, 347–352. http://link.springer.com/article/10.1023/A:1026701028901.

Bullock JM, Chapman D, Schafer S, Roy D, Haynes T, Beal S, Tinch R (2012). Assessing and controlling the spread and the effects of common ragweed in Europe (ENV.B2/ETU/2010/0037).

Caillaud D, Thibaudon M, Martin S, Segala C, Besancenot JP, Clot B, & Francois H (2014). Short-term effects of airborne ragweed pollen on clinical symptoms of hay fever in a panel of 30 patients. Retrieved from Journal of Investigational Allergology and Clinical Immunology, 24(4), 249–256. http://www.ncbi.nlm.nih.gov/pubmed/25219107. [PubMed: 25219107]

Chance CM, Coops NC, Plowright AA, Tooker TR, Christen A, & Aven N (2016). Invasive shrub mapping in an urban environment from hyperspectral and LiDAR-derived attributes. Frontiers in Plant Science, 07, 1–19. 10.3389/fpls.2016.01528..

Chapman DS, Haynes T, Beal S, Essl F, & Bullock JM (2014). Phenology predicts the native and invasive range limits of common ragweed. Global Change Biology, 20(1), 192–202. 10.1111/gcb.12380. [PubMed: 24038855]

Charalampopoulos A, Lazarina M, Tsiripidis I, & Vokou D (2018). Quantifying the relationship between airborne pollen and vegetation in the urban environment. Aerobiologia, 34(3), 1–16. 10.1007/s10453-018-9513-y.

Csornai G, Mikus G, Nador G, Hubik I, Laszlo I, & Suba Z (2009). Integration of hightech components for operating ragweed mapping and control system in Hungary using remote sensing and GIS. In Gartner G, & Ortag F (Eds.), Cartography in central and Eastern Europe (pp. 405–415). Berlin: Springer-Verlag. 10.1007/978-3-642-03294-3_25.

Cunze S, Leiblein MC, & Tackenberg O (2013). Range expansion of *Ambrosia artemisiifolia* in Europe is promoted by climate change. ISRN Ecology, 2013, 1–9. 10.1155/2013/610126.

Darrow L, Hess J, Rogers C, Tolbert PE, Klein M, & Sarnat SE (2012). Ambient pollen concentrations and emergency department visits for asthma and wheeze. The Journal of Allergy and Clinical Immunology, 130(3), 630–638.e4. 10.1016/j.jaci.2012.06.020. [PubMed: 22840851]

Denwood MJ (2016). runjags: An R package providing interface utilities, model templates, parallel computing methods and additional distributions for MCMC models in JAGS. Journal of Statistical Software, 71(9)10.18637/jss.v071.i09.

*Landsc Urban Plan. Author manuscript; available in PMC 2020 October 01.*
Dullinger S, Kleinbauer I, Peterseil J, Smolik M, & Essl F (2009). Niche based distribution modelling of an invasive alien plant: Effects of population status, pro-pagule pressure and invasion history. Biological Invasions, 11(10), 2401–2414. https://doi.org/10.1007/s10530-009-9424-5.

Essl F, Biró K, Brandes D, Broennimann O, Bullock JM, Chapman DS, … Follak S. (2015). Biological Flora of the British Isles: Ambrosia artemisiifolia. Journal of Ecology, 103(4), 1069–1098. 10.1111/1365-2745.12424.

Fawcett T (2006). An introduction to ROC analysis. Pattern Recognition Letters, 27(8), 861–874. 10.1016/j.patrec.2005.10.010.

Gelman A, & Hill J (2007). Data analysis using regression and multilevel/hierarchical models. Cambridge, UK: Cambridge University Press.

Gleason JA, Bielory L, & Fagliano JA (2014). Associations between ozone, PM2.5, and four pollen types on emergency department pediatric asthma events during the warm season in New Jersey: A case-crossover study. Environmental Research, 132, 421–429. 10.1016/j.envres.2014.03.035. [PubMed: 24858282]

Grau J, Grosse I, & Keilwagen J (2015). PRROC: Computing and visualizing Precision-recall and receiver operating characteristic curves in R. Bioinformatics, 31(15), 2595–2597. 10.1093/bioinformatics/btv153. [PubMed: 25810428]

He KS, Bradley BA, Cord AF, Rocchini D, Tuanmu M-N, Schmidtlein S, Pettorelli N (2015). Will remote sensing shape the next generation of species distribution models? Remote Sensing in Ecology and Conservation, 1 (1), 4–18. 10.1002/rase2.7.

Héguy L, Garneau M, Goldberg MS, Raphoz M, Guay F, & Valois MF (2008). Associations between grass and weed pollen and emergency department visits for asthma among children in Montreal. Environmental Research, 106(2), 203–211. 10.1016/j.envres.2007.10.005. [PubMed: 18093580]

Hjort J, Hugg TT, Antikainen H, Rusanen J, Sofiev M, Jaakkola MS, & Jaakkola JJK (2015). Fine-scale exposure to allergenic pollen in the urban environment: Evaluation of land use regression approach. Environmental Health Perspectives, 124(5), 619–626. 10.1289/ehp.1509761. [PubMed: 26452296]

Karrer G, Skjøth CA, Sikoparija B, Smith M, Berger U, & Essl F (2015). Ragweed (Ambrosia) pollen source inventory for Austria. Science of the Total Environment, 523, 120–128. 10.1016/j.scitotenv.2015.03.108.

Katz DSW, & Carey TS (2014). Heterogeneity in ragweed pollen exposure is determined by plant composition at small spatial scales. Science of The Total Environment, 485, 435–440. 10.1016/j.scitotenv.2014.03.099.

Katz DSW, Connor Barrie BT, & Carey TS (2014). Urban ragweed populations in vacant lots: An ecological perspective on management. Urban Forestry and Urban Greening 13, 756–760. 10.1016/j.ufug.2014.06.001.

Katz DSW, Dzul A, Kendel A, & Batterman SA (2019). Effect of intra-urban temperature variation on tree flowering phenology, airborne pollen, and measurement error in epidemiological studies of allergic pollen. Science of the Total Environment, 653, 1213–1222. 10.1016/j.scitotenv.2018.11.020.

Khosravipour A, Skidmore AK, Isenburg M, Wang T, & Hussin YA (2014). Generating pit-free canopy height models from airborne LiDAR. Photogrammetric Engineering and Remote Sensing, 80(9), 863–872. 10.14358/PERS.80.9.863.

Lake IR, Jones NR, Agnew M, Goodess CM, Giorgi F, Hamaoui-Laguel L., Epstein MM. (2016). Climate change and future pollen allergy in Europe. Environmental Health Perspectives, 125(3), 385–391. 10.1289/EHP173. [PubMed: 27557093]

Leiblein-Wild MC, Steinkamp J, Hickler T, & Tackenberg O (2016). Modelling the potential distribution, net primary production and phenology of common ragweed with a physiological model. Journal of Biogeography, 43(3), 544–554. 10.1111/jbi.12646.

Leutner B, & Horning N (2018). RStoolBox: Tools for remote sensing data analysis. Retrieved from https://cran.r-project.org/web/packages/RStoolbox/index.html.

Lovasi GS, O’Neil-Dunne JPM, Lu JWT, Sheehan D, Perzanowski MS, Macfaden SW, Rundle A (2013). Urban tree canopy and asthma, wheeze, rhinitis, and allergic sensitization to tree pollen...
in a New York city birth cohort. Environmental Health Perspectives, 121(4), 494–500. 10.1289/ehp.1205513. [PubMed: 23322788]

Makra L, Matyasovszky I, Bálint B, & Csépe Z (2014). Association of allergic rhinitis or asthma with pollen and chemical pollutants in Szeged, Hungary, 1999-2007. International Journal of Biometeorology, 58(5), 753–768. 10.1007/S00484-013-0656-9. [PubMed: 23558448]

Martin MD, Chamecki M, Brush GS, Meneveau C, & Parlange MB (2009). Pollen clumping and wind dispersal in an invasive angiosperm. American Journal of Botany, 96(9), 1703–1711. 10.3732/ajb.0800407. [PubMed: 21622356]

Meya-Manzano JM, Sadyś M, Tormo-Molina R, Fernández-Rodríguez S, Oteros J, Silva-Palacios I, & Gonzalo-Garjo A (2017). Relationships between airborne pollen grains, wind direction and land cover using GIS and circular statistics. Science of The Total Environment, https://doi.org/10.1016/j.scitotenv.2017.01.085.

Motreff Y, Golliot F, Calleja M, Le Pape A, Fuhrman C, Ferrara I, & Plaisant I (2014). Short-term effect of pollen exposure on drug consumption for allergic rhinitis and conjunctivitis. Aerobiologia, 30(1), 35–44. 10.1007/s10453-013-9307-1.

Nowak M, Szymanska A, & Grewling L (2012). Allergic risk zones of plane tree pollen (Platanus sp.) in Poznan. Postepy Dermatologii i Alergologii, 29(3), 156–160.

O’Neil-Dunne J, MacFaden S, & Royar A (2014). A versatile, production-oriented approach to high-resolution tree-canopy mapping in urban and suburban landscapes using GEObIA and data fusion. Remote Sensing, 6(12), 12837–12865. 10.3390/rs61212837.

Pickett S, & Baskin J (1973). The role of temperature and light in the germination behavior of Ambrosia artemisiifolia. Retrieved from Bulletin of the Torrey Botanical Club, 100(3), 165–170.

http://www.jstor.org/stable/10.2307/2484628.

Plummer M (2003). JAGS: A program for analysis of Bayesian graphical models using Gibbs sampling. In Proceedings of the 3rd International Workshop on Distributed Statistical Computing (pp. 1–10). Vienna, Austria. Retrieved from http://www.ci.tuwien.ac.at/Conferences/DSC-2003/Drafts/Plummer.pdf.

Quest Diagnostics Health Trends. (2011). The largest study of allergy testing in the United States. Retrieved from http://www.questdiagnostics.com/dms/Documents/Other/2011_QD_AllergyReport.pdf.

Raynor GS, & Ogden EC (1965). Twenty-four-hour dispersion of ragweed pollen from known sources. Technical Report from the Brookhaven National Library, No. BNL 9. Retrieved from physical ecology.

Raynor, G, Ogden, E,… Hayes J. (1970). Dispersion and deposition of ragweed pollen from experimental sources. Retrieved from Journal of Applied Meteorology, 9, 885–895. http://adsabs.harvard.edu/abs/970JApMe…9.885R.

Raynor Gilbert S., Smith ME, Singer IA, & Ogden EC. (1961). Pollen sampling and dispersion studies at Brookhaven national laboratory. Journal of the Air Pollution Control Association, 11(12), 557–584. 10.1080/00022470.1961.10468037. [PubMed: 14490748]

Roussel J, & Auty D (2017). lidR: Airborne LiDAR data manipulation and visualization for forestry applications. Retrieved from https://github.com/Jean-Romain/lidR.

Sallam MF, Fizer C, Pliant AN, & Whung PY (2017). Systematic review: Land cover, meteorological, and socioeconomic determinants of aedes mosquito habitat for risk mapping. International Journal of Environmental Research and Public Health, 14(10), 10.3390/ijerph14101230.

Salo P, Calatroni A, Gergen P, Hoppin J, Sever M, Jaramillo R, … Zeldin D. (2011). Allergy-related outcomes in relation to serum IgE: Results from the National Health and Nutrition Examination Survey 2005-2006. Journal of Allergy and Clinical Immunology, 127(5), 1226–1235. https://doi.org/10.1016/j.jaci.2010.12.1106.Allergy-related.

Simard M, & Benoit D (2011). Effect of repetitive moving on common ragweed (Ambrosia artemisiifolia L.) pollen and seed production. Retrieved from www.ncbi.nlm.nih.gov/pubmed/21736270 Annals of Agricultural and Environmental Medicine, 18(1), 55–62. [PubMed: 21736270]
Skjøth C, Ørby PV, Becker T, Geels C, Sehlüssen V, Sigsgaard T, & Hertel O. (2013). Identifying urban sources as cause of elevated grass pollen concentrations using GIS and remote sensing. Biogeosciences, 10(1), 541–554. 10.5194/bg-10-541-2013.

Smith M, Cecchi L, Skjøth CA, Karrer G, & Šikoparija B (2013). Common ragweed: A threat to environmental health in Europe. Environment International, 61, 115–126. https://doi.org/10.1016/j.envint.2013.08.005. [PubMed: 24140540]

Thibaudon M, Šikoparija B, Oliver G, Smith M, & Skjøth CA (2014). Ragweed pollen source inventory for France - The second largest centre of Ambrosia in Europe. Atmospheric Environment, 83, 62–71. 10.1016/j.atmosenv.2013.10.057.

Tuanmu MN, Vina A, Bearer S, Xu W, Ouyang Z, Zhang H, & Liu J (2010). Mapping understory vegetation using phenological characteristics derived from remotely sensed data. Remote Sensing of Environment, 114(8), 1833–1844. 10.1016/j.rse.2010.03.008.

Urbanowdcz C, Hutrya LR, & Stinson KA (2018). The effects of urbanization and land use on ragweed distribution. Ecosphere, 9(12), 1–7. 10.1002/ecs2.2512.

US Census Bureau. (2016). 2012-2016 American Community Survey 5-year estimates. Retrieved from https://www.census.gov/geo/maps-data/data/tiger-data.html.

Vincent G, & Bergeron Y (1985). Weed synecology and dynamics in urban environment. Retrieved from Urban Ecology, 9, 161–175. http://www.sciencedirect.com/science/article/pii/030440098590004X.

Wang O, Zachmann LJ, Sesnie SE, Olsson AD, & Dickson BG (2014). An iterative and targeted sampling design informed by habitat suitability models for detecting focal plant species over extensive areas. PLoS One, 9(7), 10.1371/journal.pone.0101196.

Weinberger KR, Kinney PL, & Lovasi GS (2015). A review of spatial variation of allergenic tree pollen within cities. Arboriculture and Urban Forestry, 41(2), 57–68.

Weinberger KR, Kinney PL, Robinson GS, Sheehan D, Kheirbek I, Matte TD, & Lovasi GS (2018). Levels and determinants of tree pollen in New York City. Journal of Exposure Science and Environmental Epidemiology, 28(2), 119–124. 10.1038/jes.2016.72. [PubMed: 28000684]

Werchan B, Werchan M, Mücke H-G, Gauger U, Simoleit A, Zuberbier T, & Bergmann K-C (2017). Spatial distribution of allergenic pollen through a large metropolitan area. Environmental Monitoring and Assessment, 189(4), 169. 10.1007/s10661-017-5876-8. [PubMed: 28316024]

Zhong W, Levin L, Reponen T, Hershey GK, Adhikari A, Shukla R, & Lemasters G (2006). Analysis of short-term influences of ambient aeroallergens on pediatric asthma hospital visits. Science of the Total Environment, 370, 330–336. 10.1016/j.scitotenv.2006.06.019.

Zink K, Kaufmann P, Petitpierre B, Broennimann O, Guisan A, Gentilini E, & Rotach MW (2016). Numerical ragweed pollen forecasts using different source maps: A comparison for France. International Journal of Biometeorology, 1–11. 10.1007/s00484-016-1188-x.
Fig. 1.
A) Ragweed transects (colored lines) within the City of Detroit (black outline); B) Example of part of a transect (solid line) in a residential area including vacant lots and searched area (shaded area); C) Common ragweed in a vacant lot.
Fig. 2.
Workflow diagram for ragweed habitat suitability model and ragweed pollen production estimates.
Fig. 3.
Predicted ragweed presence rates in several common land cover, land usage, demolition ("demo"), and spectral reflectance categories in 2016 and 2017. All classes besides buildings and trees are for ground pixels; other values are set at 0. For the image classifications categories values are set at 100% cover.
Fig. 4.
A) Predicted ragweed presence across the City of Detroit; B) Predicted ragweed occupancy in a small section of the city; C) Aerial imagery of the inset area (note the recent demolitions [light brown swaths] and correspondingly higher probability of ragweed presence there).
Fig. 5.
Maps of predicted ragweed pollen production (pollen per m$^2$) across Detroit in 2016 and 2017 averaged at various radii around each cell.
Table 1

Variables included in the final model and associated parameter estimates. Variables whose 95% credible intervals do not overlap zero are considered statistically significant.

| Data type             | Variable                     | Parameter | Mean   | SD    | 95% CI         |
|-----------------------|------------------------------|-----------|--------|-------|----------------|
| Land cover            | Trees                        | $\beta_1$ | -6.16  | 0.42  | -7.02 to -5.37 |
|                       | Buildings                    | $\beta_2$ | -83.45 | 58.67 | -223.56 to -9.51 |
|                       | Ground                       | $\beta_3$ | -4.89  | 0.30  | -5.50 to -4.35 |
| Land use & social data| Vacant lot                   | $\delta_1$ | 1.38   | 0.27  | 0.88-1.92       |
|                       | Year (2018)                  | $\delta_2$ | -1.61  | 0.37  | -2.38 to -0.94 |
|                       | Demolition (within 16 m)     | $\delta_3$ | 0.98   | 0.26  | 0.47-1.49       |
|                       | Road                         | $\delta_4$ | -1.76  | 0.65  | -3.17 to -0.64 |
|                       | Demolition (within 64 m)     | $\delta_5$ | 0.73   | 0.27  | 0.19-1.28       |
| Imagery               | Class 29 (thin vegetation)   | $\alpha_1$ | 2.53   | 0.90  | 0.68-4.21       |
|                       | Class 32 (light pavement)    | $\alpha_2$ | -9.21  | 4.75  | -20.59 to -2.18 |
|                       | Class 44 (dark pavement)     | $\alpha_3$ | -46.01 | 29.92 | -118.89 to -6.27 |
|                       | Class 5 (lawn)               | $\alpha_4$ | -4.06  | 2.17  | -8.82 to -0.51 |