scientific reports

Check for updates

OPEN

Cannabis pollen dispersal across the United States

Manu Nimmala[1]✉, Shane D. Ross[2,4] & Hosein Foroutan[3,4]

For the recently legalized US hemp industry (*Cannabis sativa*), cross-pollination between neighboring fields has become a significant challenge, leading to contaminated seeds, reduced oil yields, and in some cases, mandated crop destruction. As a step towards assessing hemp cross-pollination risk, this study characterizes the seasonal and spatial patterns in windborne hemp pollen dispersal spanning the conterminous United States (CONUS). By leveraging meteorological data obtained through mesoscale model simulations, we have driven Lagrangian Stochastic models to simulate wind-borne hemp pollen dispersion across CONUS on a county-by-county basis for five months from July to November, encompassing the potential flowering season for industrial hemp. Our findings reveal that pollen deposition rates escalate from summer to autumn due to the reduction in convective activity during daytime and the increase in wind shear at night as the season progresses. We find diurnal variations in pollen dispersion: nighttime conditions favor deposition in proximity to the source, while daytime conditions facilitate broader dispersal albeit with reduced deposition rates. These shifting weather patterns give rise to specific regions of CONUS more vulnerable to hemp cross-pollination.

The 2014 and 2018 US Farm Bills legalized the production of industrial hemp (*Cannabis sativa*) for cannabidiols, seed, and fiber[1]. This nascent industry has been challenged by wind-blown cross-pollination between neighboring hemp fields, leading to contaminated seeds, reduced oil yields, and in some cases, mandated crop destruction[2,3]. Financial impacts reported in a 2022 Colorado survey[2] ranged from $12,000 to millions of dollars, with an Oregon lawsuit alleging damages of over $8 million[4]. Economic modeling[3] shows that the industry will transition away from cannabidiol hemp production entirely without effective cross-pollination mitigation strategies.

As hemp production has only recently been legalized[1], there is a deficit in hemp dispersal research. The only study quantifying hemp pollen dispersal as a function of distance from a known source is an experiment by Small and Antle[5]. They sampled hemp pollen for three weeks at distances of up to 400 meters from a source field and observed significant deposition even at the edge of their domain, 17,000 pollen grains/m^2/day, enough to "achieve excellent seed set", i.e., successfully cross-pollinate. The authors noted that due to its small size ($\sim$30 microns) hemp pollen travels farther and deposits in greater quantities than other wind-pollinated crops, and that it is prolific—each male flower can release up to 350,000 pollen grains, and there are potentially hundreds of flowers on larger plants[6]. A single male plant can therefore release about 100 million pollen grains. Recommended isolation distances are far greater than the experimental domain, typically varying between 1 and 5 km[7,8], but there have been reports of cross-pollination up to 20 km[9] and even 48–96 km away[2]. Two back-trajectory studies have demonstrated that *Cannabis* pollen likely travelled over 200 km, from Northern Africa to Spain[10,11]. This indicates that hemp pollen has great potential for long-distance transport, and that the 'fat tail' of the hemp pollen dispersal kernel could play an outsized role in cross-pollination between fields.

Dispersal modeling studies show that the fat tail in wind-borne dispersal is highly sensitive to changes in meteorological conditions, particularly the combined effects of shear and convective turbulence. During the day, solar heating of the surface induces a positive heat flux that creates large-scale convective updrafts. Shear-driven turbulence arises as horizontal wind passes over rough surfaces. One study found that rising temperatures, correlated with increasing heat flux, led to a greater proportion of seeds traveling beyond 100 meters in simulations[12]. Another found that sustained updrafts caused dandelion seeds to disperse further, while horizontal wind speed did not play a factor[13]. In contrast, Soons et al.[14] found that horizontal wind velocity was the primary driver of downwind transport, and heat flux only played a role when wind velocity was low (< 4 m/s). Understanding such patterns in variation of the tail would help inform cross-pollination mitigation strategies.

Two dispersal modeling studies have identified seasonal and diurnal patterns in the variation of wind-borne dispersal kernels. Oneto et al.[15] used the Hybrid Single-Particle Lagrangian Integrated Trajectory (HYSPLIT)

[1]Engineering Science and Mechanics, Virginia Tech, Blacksburg, VA 24061, USA. [2]Aerospace and Ocean Engineering, Virginia Tech, Blacksburg, VA 24061, USA. [3]Civil and Environmental Engineering, Virginia Tech, Blacksburg, VA 24061, USA. [4]These authors contributed equally: Shane D. Ross and Hosein Foroutan. ✉email: nimmala@vt.edu

model to simulate fungal spores released at ten North American locations in January, April, July, and October, 2014. They found a strong diurnal pattern in average flight times, with spores staying in the air longer during the day than at night. They also observed seasonal changes, with the longest flight times in July and lowest in January. Savage et al.[16] simulated spore dispersal using hourly meteorological inputs from a large-scale weather model at two towns in Western Australia for June and September 2007, early winter and early spring, respectively. They found seasonal and diurnal changes in the number of spores travelling past 10 km, and differences between the two towns, aligning with seasonal and diurnal changes in temperature and wind velocity. These studies suggest contiguous spatial patterns in dispersal on a country-wide scale.

In this study, we seek seasonal and spatial patterns in pollen dispersal spanning the conterminous United States (CONUS), revealing regions more prone to cross-pollination. We extend the methodology of Savage et al.[16], using meteorological data provided by a mesoscale model simulation to drive Lagrangian Stochastic (LS) models of pollen dispersion for each county in the United States over five months. The LS model is ideal for examining the sensitivity of dispersal due to shear and convection, as it more naturally captures the variations of turbulent flow using stochasticity. It is an application of Brownian motion to turbulent diffusion, in which the trajectories of many particles through the air are modeled as random walks. By releasing thousands of particles and computing an ensemble average of their trajectories, we can determine the relative concentration at any point in the domain and the mean shape of the plume. Therefore, they require a fraction of the computational resources of more resolved Eulerian models like Large Eddy Simulations. Although conventional Gaussian plume models are computationally lighter than LS models, their treatment of turbulence is more prescribed. Modifications have been made to incorporate effects like convection in Gaussian plume models (for example, the AERMOD model[17]), but these require more parameters and increase complexity[18].

We used two LS model formulations: a convective boundary layer model[19–21] for unstable (typically day) conditions and a surface layer model[22] for stable (night) conditions. To drive the LS model, we used meteorological fields obtained from a Weather Research and Forecasting (WRF) model simulation over CONUS for the entire year of 2016[23]. This high-resolution meteorological dataset, developed by the U.S. Environmental Protection Agency to support modeling applications, comprises an hourly time series of weather conditions on a 12 km-square horizontal grid and has been extensively validated[24]. For each county, we extracted the weather data at the grid point nearest to its centroid and averaged across local noon and midnight hours for each month from July to November, to represent average "day" and "night" conditions respectively. We performed LS simulations for day and night conditions, for five months from July to November, for each of 3,107 counties in the CONUS, totalling to 31,070 simulations. In this study, we used 2D LS models, in which we simulate pollen travelling in the downwind and vertical directions. From each simulation, we compute a dispersal kernel by counting the number of particles which have deposited in the simulation domain within 250 meter-wide bins up to 50 km downwind of the source. The meteorological conditions are assumed to be statistically stationary and horizontally homogeneous for each simulation.

To the best of our knowledge, this is the first simulation study of hemp pollen dispersal. It is also the first large-scale simulation study of the inhomogeneity of pollen dispersal across regions and seasons.

Results and discussion

Simulation of day and night pollen dispersion over five months reveals significant seasonal and spatial variations, particularly in the tail of the dispersal kernel. Each simulation yielded a dispersal kernel, or number of particles deposited downwind from the source in 250 m wide bins, normalized by the number of particles released. Figure 1a,b show median day and night dispersal kernels on a log scale by month for each of nine US climate divisions[25], in order to compare between climatically different regions. We observe depositions up to 50 km downwind, the edge of our domain, which is the limit of applicability of our LS model.

The tail of the dispersal kernel varies seasonally and spatially

Simulations of day and night pollen dispersion over five months yields variation only in the tail of the dispersal kernel. For all climate regions, in both day and night conditions, Fig. 1 shows a steep decline in depositions by two orders of magnitude within the first few kilometers of the source. Approximately 70% of simulated pollen is deposited in the first bin alone for all cases. Figure 2a shows that across all simulations, dispersal kernels decreased to 1% of released particles within 3 km of the source. Although there is a slight increase in distance for nighttime conditions, this region of steep decline is indistinguishable across counties regardless of region and seasonal weather changes.

While this steep decline in depositions appears to support commonly-used hemp isolation distances (< 5 km[7,8]), even 1% of 100 million pollen grains would result in 1 million pollen depositing at that distance. In Fig. 2b, lowering the threshold to 0.1% of released particles results in far more spread, 1–10 km during the day, and 10–15 km at night. Further decreasing the threshold to 0.01% results in distances varying throughout the entire domain, as shown in Fig. 2c. This fat tailed deposition kernel is common for wind-dispersed species[26,27], and poses challenges when computing the risk of rare events in the tail, e.g., burning embers from a wildfire[28] or cross-pollination. For hemp in particular, the Small and Antle experiment[5] provides evidence that even reduced depositions at the tail of the distribution can result in effective cross-pollination. Given the prolific nature of hemp pollen, potentially massive fields, and reports of hemp pollen travelling well beyond established isolation distances, the fat tail of the dispersal kernel becomes necessary to assess cross-pollination risk[27,29].

We find that the tail of the dispersal kernel below the 0.1% and 0.01% thresholds and beyond 3 km, shows considerable variability. Figure 2b,c show stark differences between day and night simulations, driven by diurnal differences in wind conditions. For more detail, see Supplementary Fig. S4. Below the 0.01% threshold, we

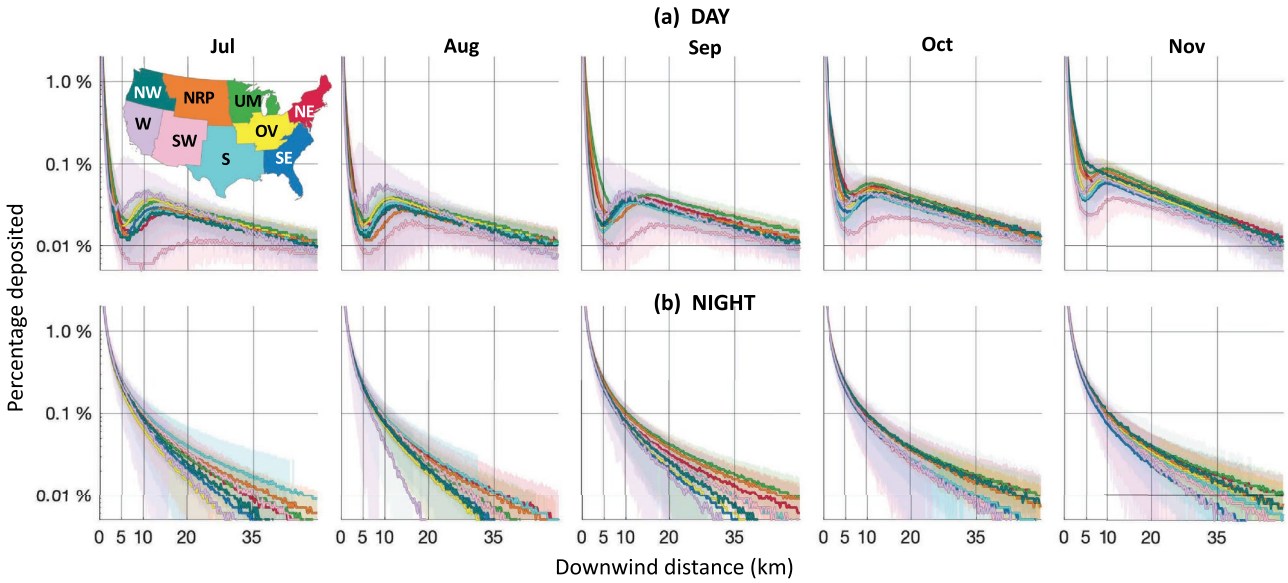


Fig. 1. Median dispersal kernels for each month during (**a**) daytime and (**b**) nighttime, separated by US climate region: Northeast (NE), Upper Midwest (UM), Ohio Valley (OV), Southeast (SE), Northern Rockies and Plains (NRP), South (S), Southwest (SW), Northwest (NW), and West (W). Dispersal kernels are formed by counting depositions within 250 meter-wide bins up to 50 km downwind of the source, normalized by the amount released. Shading represents data between the 10th and 90th percentiles. Note that the vertical axis is a log scale.

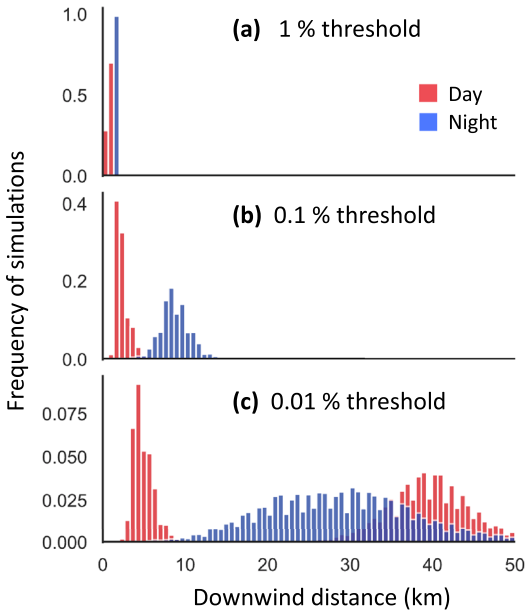


Fig. 2. Distances at which dispersal kernels first fall below a threshold: (**a**) 1%, (**b**) 0.1%, and (**c**) 0.01%. Red represents day simulations, while blue represents night.

observe a large spread in nighttime threshold distances and two peaks for day simulations, which point to large-scale regional and seasonal shifts in wind conditions.

Daytime seasonal and spatial patterns

In Fig. 1, daytime dispersal kernels for all climate regions exhibit a steady rise from July to November. This increase is responsible for the second peak in daytime 0.01% threshold distances, which is dominated by simulations later in the season. Although all regions experience increase over the season, the Southwest region maintains the least depositions throughout. In the peak summer months of July and August, the Southwest region experiences the lowest depositions, as do the Northwest, Northern Rockies & Plains, and Northeast. By October and

November these latter three regions exhibit an almost 10-fold increase, shifting from relatively low depositions to the highest, on par with the Upper Midwest and Ohio Valley.

Seasonal shifts are most apparent between 5 and 10 km downwind, where overall depositions increase by nearly an order of magnitude. At this distance, Fig. 1 shows a distinctive local minimum near the source for nearly all simulations. The daytime dispersal dip in an otherwise monotonically decreasing curve is due to updrafts from convective turbulence[30,31], and can be interpreted as a region of relatively less deposition, or a "pollen shadow", in the near-field downwind of the source. Beyond the pollen shadow, there is relatively less seasonal and regional variation in depositions, indicating that in daytime, these downwind distances are not as strongly tied to patterns in underlying meteorological parameters.

Mapping out daytime deposition values in Fig. 4a at 5 km, 10 km, 20 km, and 35 km downwind reveals contiguous, large-scale seasonal and spatial patterns. Within the pollen shadow, at 5 km downwind, Northern counties are the first to experience increases in deposition. From September, we see a region of higher depositions in California and the Upper Midwest. That region extends to the northernmost counties by October, coalescing into a band above about 40° N latitude in November. Further downwind, beyond the pollen shadow, this pattern of northern seasonal increase is not as apparent; only the Southwest stands out with the lowest depositions throughout the season.

We observe the lowest depositions in simulations with higher boundary layer height, z_i, and greater convective velocity, w^*. High w^* and z_i together indicate greater buoyancy associated with the surface heat flux and more convective turbulence[32]. Scatter plots and correlation values between daytime depositions and these meteorological parameters are provided in Supplementary Fig. S2 and the monthly heatmaps are shown in Supplementary Fig. S6. High convective conditions in summer lead to more pollen uplifting and less deposition, particularly in the pollen shadow. More pollen is uplifted, carried far from the source, before descending in small quantities at great distances. A reduction in convective conditions from summer to fall explains the pattern of deposition increase for northern regions, particularly within the pollen shadow. It is also why the Southwest exhibits low depositions throughout the season. Greater convective conditions make long-distance transport of pollen more likely[12,15,16], but results in fewer depositions within the domain.

Our results align with other dispersal studies, which show that greater sensible heat flux and warming temperatures during the day led to greater transport distances[12,13,16], particularly in combination with increased wind speed[14]. In our results, however, neither the 10-m wind speed (estimated roughly as $10u^*$) nor the Monin-Obukhov length, L, influenced deposition counts, indicating that shear-driven turbulence did not play a major role in daytime dispersal patterns. This could be due to the monthly averaging of the meteorological input parameters. For example, monthly-averaged u^* only varied between 0.45 and 0.65 m/s, or maximum variations in 10-m wind speed of 2 m/s. It is likely that averaging resulted in less variation, allowing convective conditions to govern deposition patterns within the domain.

In summary, during the day, we identify large-scale contiguous spatial patterns that shift from summer to fall. The Southwest maintains the lowest depositions throughout the season because it experiences greater convective conditions than all other regions. On the other hand, northern counties shifted from comparatively low to high depositions relative to other climate regions due to a decrease in convective conditions in the fall months. This is consistent with typical CONUS weather trends; Northern climate regions experience changing seasons more strongly, and daytime dispersal is particularly dependent on these seasonal factors.

Nighttime seasonal and spatial patterns

Unlike the daytime curves, night-time dispersal kernels for each month show a monotonic decrease with downwind distance, as shown in Fig. 1b. Within the first 10 km, depositions at night are ten times greater than during the day. Relative to these large values, spatial patterns and seasonal differences only become clear beyond about 10 km. Beyond this distance, we observe slight overall increase in deposition primarily in October and November.

While we do not see a major seasonal increase at night, shifting spatial patterns are discernible in both the heat maps and dispersion kernels. Figure 4b shows night-time depositions by county at 5, 10, 20, and 35 km downwind of the source. Observing heatmaps at 10 km and beyond, in July and August, there is a swathe of high depositions in the center of the country, beginning with the South region and extending into the Northern Rockies & Plains (NRP). By September, the South region is no longer as prominent, and by October, the swathe of high depositions has extended into the Upper Midwest (UM) and NRP. The Northeast (NE) region also progressively increases in depositions over the season. By November, the regions with the greatest deposition include the UM, NRP, and NE, while the least deposition occur in the Southeast and West regions.

We find that regions of least deposition correspond to high friction velocity, u^*, high boundary layer height, z_i, lower roughness length, z_0, and high Monin-Obukhov length $|L|$. Scatter plots and correlation values between night-time depositions and these meteorological parameters are provided in Supplementary Fig. S3. These parameters indicate more neutral conditions and greater wind shear, resulting in pollen travelling further from the source and depositing in greater amounts[14]. Our results show that greater u^*, i.e., greater horizontal wind speed, is primarily responsible for variations in night time dispersal, and the slight increase in depositions in the cooler months of October and November. This aligns with previous dispersal studies, which show that particles travel further[16] and remain airborne for longer[15] in winter than in summer months.

Overall, we find that night-time dispersal kernels are dictated by wind speed, or shear-driven turbulence. This results in more depositions further downwind in cooler months, where depositions increased with greater wind speeds.

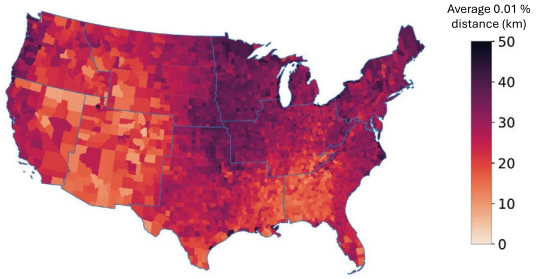


Fig. 3. Heat map of 0.01%-distances averaged over all day and night simulations from July to November for each county.

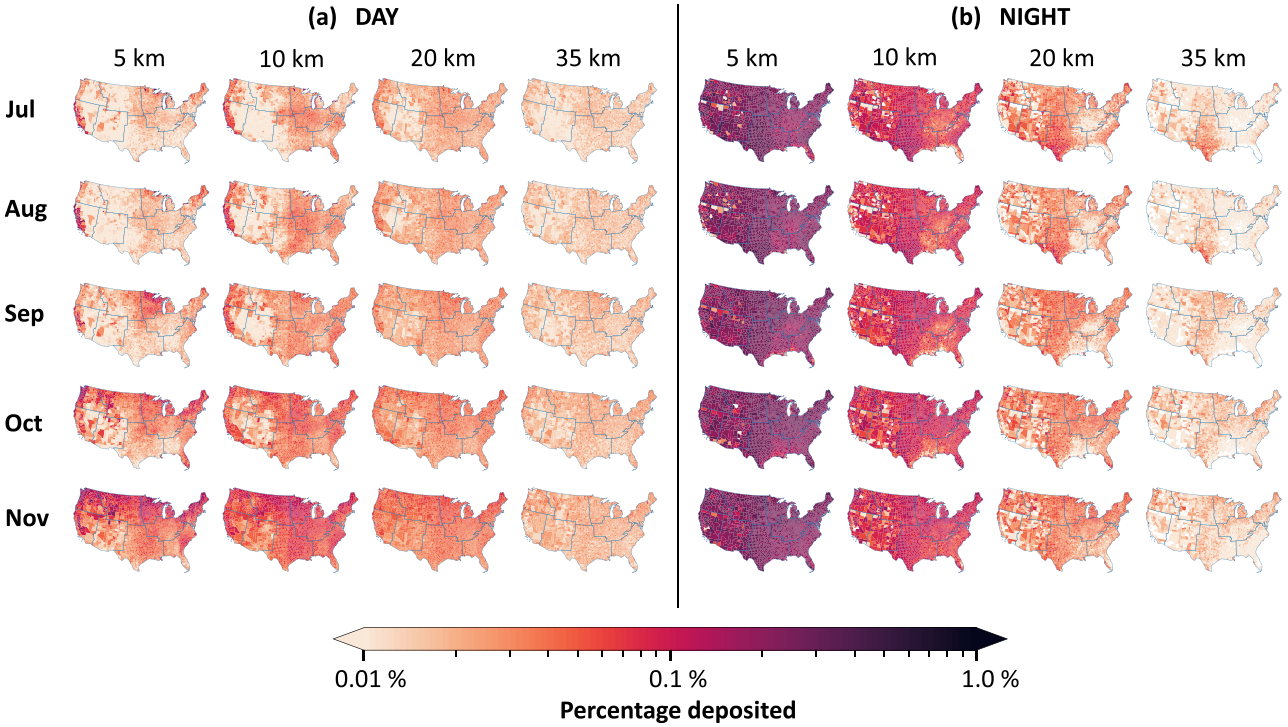


Fig. 4. Percent of particles deposited in 250 m-wide bins at downwind distances of 5, 10, 25, and 35 km for each county: (**a**) daytime simulations, (**b**) nighttime simulations. Note that the colorbar is a log scale.

Reconciling day and night patterns

We observe strong diurnal patterns and find that night-time dispersal dominates consideration of cross-pollination risk near the source. Within approximately 20 km of the source, night-time depositions are one to two orders of magnitude greater than during the day, as shown in Figs. 1 and 4. Nearly all released particles are deposited by 20 km at night—an average of 97% across night-time cases, compared to only 81% during the day. Cumulative depositions are shown in Supplementary Fig. S5. This results in a stark difference in cross-pollination risk between day and night, showing that nighttime dispersal is more important to consider within the domain and within 20 km.

Beyond this distance, nighttime dispersal kernels experience a steep decline in depositions, while daytime kernels possess a fatter tail. We can see this at 35 km in Fig. 1, where the daytime kernels have a shallower slope than and in Fig. 4, where most regions during the day are greater than at night. At night, almost all pollen is deposited near the source, but convective uplifting during the day allows for pollen to deposit in low quantities at the furthest reaches of the domain and even beyond it. Oneto et al.[15] found that spores released during the day had much longer flight times than at night, on the order of several days rather than a few hours and escaped into the stratosphere in greater numbers, while spores at night had flight times on the order of hours. For longer day flight times, pollen viability may become a factor for risk of cross-pollination[15]. Choudhary et al. found that viability of *Cannabis* pollen only decreased substantially three days after release from the anther[33]. In our study, we are only considering dispersal within 50 km of the source. Even with a slow wind speed of 1 m/s, it would only take a pollen grain 14 h to traverse the 50 km domain, thus viability need not be taken into account. Within

the domain, viability has little impact on cross-pollination risk, and so daytime dispersal patterns impact risk at the furthest reaches of our domain.

It is possible that hemp pollen only disperses during the day, as is common for many wind-dispersed species[27]. One study observed that male *Cannabis* anthers open and release pollen in the morning hours[33]. However, *Cannabis* pollen measurement studies found only slight diurnal changes in concentration[10,34], indicating that *Cannabis* pollen remains in the air throughout the day. As *Cannabis* production has only recently been legalized, there is minimal research on the diurnal timings of *Cannabis* pollen release. For these reasons, we consider both day and night dispersal in this study for risk assessment.

Cross-pollination vulnerability

While we cannot directly estimate risk of cross-pollination, as these are 2D models that do not take into account lateral spread, we can evaluate counties based on total counts of particles reaching certain distances downwind. In Fig. 3, we plot the 0.01%-distances averaged over all day and night simulations from July to November for each county as a heat map. This figure shows that across all months and time periods, the Upper Midwest, Ohio Valley, and Northeast regions have the greatest average 0.01% threshold distances—they experience the most depositions at the farthest distances. Thus, according to simulation results alone, these regions are most vulnerable to cross pollination.

However, when county-specific information such as hemp acreage and land area are incorporated, vulnerability does not necessarily reflect the same contiguous spatial patterns demonstrated in Fig. 3. In Eq. (1) below, we incorporate this information to compute a novel, dimensionless "vulnerability" metric for each county. We first normalize the dispersion area, A_{disp}, i.e., the area of a circle with radius equal to the average 0.01% threshold distance, by the land area of each county, A_{land}. This yields the fraction of a county that falls within its theoretical area of risk. We then normalize the number of acres of planted hemp in 2023 per county[35], A_{hemp}, by the land area of each county, A_{land}. This yields the proportion of land used for hemp cultivation for each county. See Supplementary Fig. S7 for heat maps of the components of the vulnerability metric. We then multiply these two factors to produce a rudimentary measure of how vulnerable a county is to cross-pollination,

$$\text{Vulnerability} = \frac{A_{\text{disp}}}{A_{\text{land}}} \times \frac{A_{\text{hemp}}}{A_{\text{land}}}. \tag{1}$$

Figure 5 shows a heat map of the vulnerability metric for all counties with nonzero hemp acreage in 2023[35]. The five states with the most land area with vulnerability greater than 10×10^{-6}—Montana, South Dakota, Idaho, Wisconsin, and Kentucky—are enlarged to illustrate vulnerable counties in more detail.

In counties with high vulnerability, large isolation distances may not be sufficient to prevent cross-pollination, as the combination of more hemp acreage and larger 0.01% threshold distances result in a greater likelihood of pollen transport across the entire county. Instead, a more comprehensive approach is necessary. A 2022 Colorado report exploring hemp cross-pollination suggested a voluntary pinning system to track where hemp is planted in a region[2]. Rather than mandating specific isolation distances, we recommend a pinning system which includes location of outdoor plants, time of planting, and anticipated flowering dates. This could be combined with an awareness of when and where pollen transport is greatest, as demonstrated in this study, to produce a dynamic time-dependent map of high-risk areas within a county.

The dependence of the vulnerability metric on dispersal distances and meteorological conditions tends to vary by regional meteorological differences on a country-wide scale. Within a state, variation in the vulnerability metric is more dependent on hemp acreage within a county. These country-wide spatial patterns and local variations

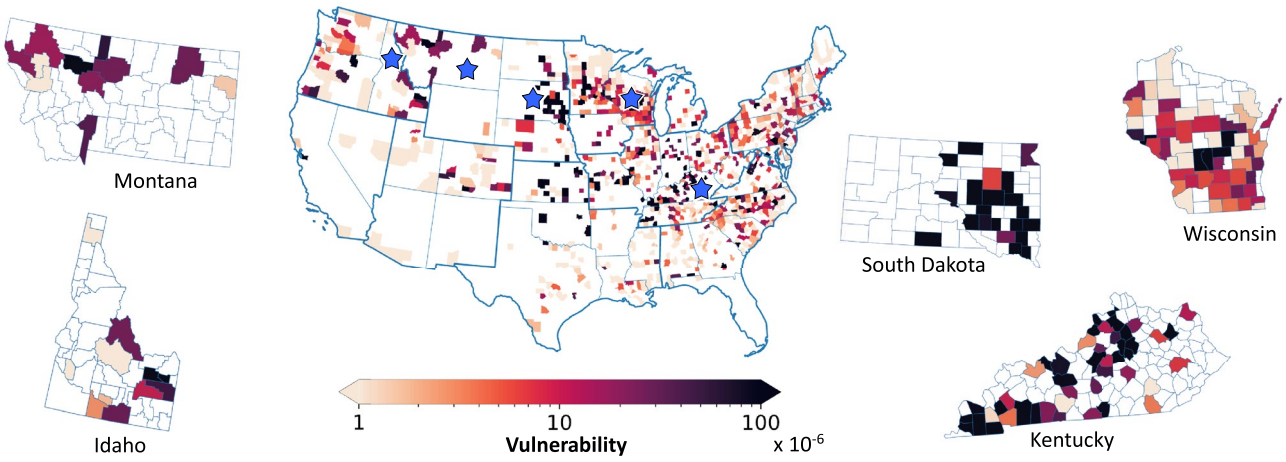


Fig. 5. Vulnerability to hemp cross-pollination across the conterminous United States. The counties with nonzero planted hemp acreage as of 2023 are shown with darker shades showing greater vulnerability. The five states with the most land area with vulnerability greater than 10×10^{-6} are shown with stars.

could be useful for potentially insuring farmers in the event of financial losses due to cross-pollination—another form of risk management, with insurance coverage and premiums varying based on region and local risk.

Weather forecasting, combined with dispersal modeling, could provide a way to predict when and where pollen will tend to travel further, rather than relying on historical weather patterns as done in this study. This would enable individual farmers to plan their crops strategically, incorporating dispersal patterns when evaluating the risks of growing one crop over another. It would also allow for voluntary community-level planning, where stakeholders make decisions together regarding when and where certain varieties should be planted in each season. Finally, local government could require sharing of crop timing and location so that more informed decisions could be made.

Cannabis is typically photosensitive, flowering as day lengths shorten below a threshold (typically 10–12 hours)[36,37]. However, this varies depending on the cultivar and planting location. A strain adapted to northern latitudes may flower in an entirely different month when planted further south[38], and there are also non-photosensitive cultivars[39]. It may be possible to strategically plan and plant crops so that flowering times between fiber/grain growers and floral hemp do not overlap. A three-year *Cannabis* pollen sampling study[34] in Tetouan, Morocco, observed that the main pollen season, when concentrations peaked, began almost a month late due to rainfalls that caused delays in planting. Strategic planting and community coordination could shift the dates of regional pollen concentration significantly. In fact, artificially reducing the day-lengths by covering crops has also previously been suggested to induce earlier flowering[40].

Strategic community planning for hemp growers would alleviate many of the challenges facing the US hemp industry today. This industry is extremely new, and is still developing the infrastructure to balance production with supply-chain capacity and consumer demand[41]. For example, in North Carolina, there was a crash in cannabidiol hemp production following a massive grower rush which exceeded demand[38]. There are also insufficient fiber processors for the state to bounce back to growing for fiber. It has been suggested that hemp grown for fiber, cannabidiols, and seed should be grown near their respective processing facilities in order to optimize production and prevent such problems[41]. These kinds of risks, in addition to the cross-pollination risk, can be managed with more intensive community planning.

Limitations and future directions

Currently, there is no single LS model that addresses both stable and unstable conditions effectively across our entire domain. Therefore, to model dispersal both during the day (typically unstable) and the night (typically stable), we chose two separate LS model formulations. Although this choice of different models for day and night might influence the observed diurnal patterns in this study, our results qualitatively align with the literature in terms of day and night differences and seasonal variation[15,16]. In addition, the LS model we use for stable conditions incorporates only shear-generated turbulence produced at the surface. In reality, turbulence in the nocturnal boundary layer is complex, involving physics such as decoupling from the surface layer, the low-level nocturnal jet, and slope effects[32,42]. Future work to identify night-time dispersal patterns might include more nuanced modeling in stable conditions. In general, more resolved, albeit more computationally expensive models, would greatly improve risk prediction. These models could incorporate more detailed physics such as release of pollen from the anthers, dispersal within a canopy, wet deposition, and even conditions specific to a farm's location like topography.

The models used in this study were shown to perform reasonably when compared to experimental results, described further in the "Methodology" section. However, we have not found previous experimental *Cannabis* pollen dispersal studies with enough information to validate the model. Experimental evidence suggests that airborne *Cannabis* pollen is ubiquitous[43], in part because of its long flight times due to its small size compared to other pollen[5]. Therefore, validation of dispersal from a known source is difficult. One approach is to use a source made of genetically engineered (GE) plants which produce pollen with fluorescent markers[44], enabling accurate source attribution. Our group is currently pursuing this in collaboration with co-workers. However, making GE *Cannabis* has proven difficult, and a study was performed instead with GE switchgrass, which produces pollen of a similar small size[45]. A paper on this combined experimental and modeling study is forthcoming.

The present study was performed using meteorological data only from 2016. This data has been validated with an extensive measurement network in the US[24], which was deemed appropriate for this proof-of-concept study. Performing this same study over multiple years could increase the robustness of our results and provide insight into possible yearly variation. For example, warming temperatures could cause changes to these seasonal and spatial patterns. Kuparinen et al.[12] demonstrated greater seed dispersal distances achieved in simulations when using increasing temperatures.

In this study, averaging meteorological data across months reduces the occurrence of extreme weather patterns and does not take into account frequency of certain conditions. Incorporating wind-direction frequency would provide directionality to cross-pollination risk assessment. For example, the Small and Antle experiment[5] measured six times more pollen deposition downwind than upwind at their source field over a period of two weeks. For future studies, a better measure of cross-pollination risk would include frequency of weather conditions and directional variability in deposition.

Furthermore, incorporating the distance between farms would provide a more sophisticated measure of county vulnerability, as was demonstrated theoretically for hemp farms in Kentucky counties[46]. Our vulnerability metric assumes one source of hemp per county, as data for the locations of individual farms are not currently available. When averaging the 0.01%-threshold distances, we weighted day and night dispersal equally, as literature describing diurnal *Cannabis* emission rates is lacking. However, including temporally varying rates of pollen emission would increase accuracy.

Conclusion

This investigation represents a pioneering effort to assess the potential risks associated with windborne hemp cross-pollination, emphasizing the variability in risk across different seasons and geographic regions. By leveraging meteorological data for an entire year, obtained through mesoscale model simulations, we have driven Lagrangian Stochastic models to simulate wind-borne pollen dispersion across the conterminous United States on a county-by-county basis. Our findings reveal that pollen deposition rates generally escalate from summer to autumn, attributed to the reduction in convective activity during daytime and the increase in wind shear at night as the season progresses. Notably, we detected pronounced diurnal variations in pollen dispersion: nighttime conditions favor deposition in proximity to the source, while daytime conditions facilitate broader dispersal albeit with reduced deposition rates. Such variability complicates the establishment of uniform isolation distances, suggesting the superiority of adaptive risk management strategies. These strategies could incorporate weather pattern considerations to mitigate cross-pollination risks more effectively and could include measures like intertemporal zoning, farm quotas, cross-pollination damage insurance, and regulatory policies.

To our knowledge, this study is unprecedented in its comprehensive simulation of pollen dispersal's regional and seasonal inhomogeneities, specifically focusing on hemp. Although this study centers on *Cannabis* pollen, the methodologies employed are broadly applicable to the dispersion of any lightweight particles. This study lays the groundwork for developing sophisticated approaches to managing agricultural cross-pollination risks, potentially influencing both policy and practice.

Methodology

Lagrangian Stochastic model formulations

For this study, we required simulation of dispersal across a wide range of wind conditions, encompassing both the convection-driven unstable conditions typical of daytime and the shear-driven stable conditions of night. There is a surface-layer LS model that has been used effectively in both conditions[22,29,47], but modeling the surface-layer alone is not sufficient in convective conditions and up to the 50 km scale we are interested in. In convective conditions in particular, we need to model the entire boundary layer, to capture both plume rise and descent. There is not currently a single LS model that addresses both conditions effectively across our entire domain. So we use two formulations: the surface-layer model (SL) for all stable conditions, and another model formulated for the convective boundary layer (CBL) for all unstable conditions.

Unstable formulation. For all unstable convective conditions, we employ a model formulated for the CBL, first introduced by Luhar et al.[19,48]. This model captures the skewed nature of the vertical wind velocity fluctuations, due to the convective updrafts and downdrafts, using the summation of two Gaussian probability distribution functions (PDFs), one representing updrafts and the other downdrafts. Luhar et al.[19] further introduced a new closure that enables the model to reduce to a single Gaussian distribution in the limit of zero skewness, typical of neutral and stable conditions, which expands the model's applicability to neutral conditions. Boehm et al.[20] adapted the model to include heavy particles, and Boehm et al.[21] introduced wind statistics profiles which merge shear-generated turbulence at the surface with convective turbulence above. Here, CBL-SL wind statistics are merged in an effort to create a smooth transition from unstable to stable regimes. Results from the original CBL model aligned well with convective fluid tank experiments[19]. Predicted concentrations from the merged model were found to reasonably compare with measured aerial pollen concentrations[21].

Stable formulation. For all stable conditions, we used the surface-layer model as described in Aylor[22]. It differs from the CBL model under neutral conditions only in that it uses a jointly Gaussian PDF in the u and w wind velocity components (downwind and vertical, respectively), resulting in better modeling at the surface. The CBL model assumes u and w wind velocity fluctuations are independent[49]. However, being a surface layer model, it incorporates only shear-generated turbulence produced at the surface. For the purpose of this study, including only the surface layer under stable conditions is sufficient, as species released in the stable boundary layer experience little vertical mixing[32]. In our simulations, pollen is released near the surface to represent release from a hemp field. Hence, we do not expect significant vertical transport above the surface layer. Results from this model have been previously compared favorably with measured pollen concentrations in stable conditions[22]. The complete model formulations for both stable and unstable conditions can be found in Supplementary Methods S1 online.

Wind statistics

LS models require wind statistics at every point in the domain, i.e., the mean, variances, covariances, and skewness. Both SL and CBL formulations assume horizontal homogeneity and stationarity, so that wind statistics vary only with height and remain constant for the duration of the simulation. Under these assumptions, we apply boundary layer scaling parameterizations to compute vertical profiles of the wind velocity statistics[21,42,50–52] as a function of five meteorological parameters: the friction velocity u^*, the Monin-Obukhov length L, the convective velocity scale w^*, the surface roughness length z_0, and boundary layer height z_i. Complete wind statistics profiles utilized in the models can be found in the Supplementary Methods S1 online.

Hemp pollen simulations

To simulate hemp pollen dispersal for each county in the CONUS, we release particles from a point source at a height of $h_0 = 2$ m. Hemp height can vary between 1-5 meters, depending on its type and growing conditions[36,53]. A study examining hemp morphology found the mean height of 16 genotypes in the 1–2 m range[54]. We varied the release height by ± 0.5 m to test the sensitivity of our results to changes in release height. We found that although increasing the magnitude of depositions changed, qualitatively, the seasonal and spatial patterns we found remained the same. This sensitivity analysis can be found in the Supplementary Fig. S1 available online.

For all simulations, we used a settling velocity of $v_s = 0.027$ m/s, based on a typical hemp pollen diameter of 30 μm[5,36], using Stokes' law. As hemp pollen is nearly spherical[36], Stokes' law provides a good approximation of settling velocity[47,55]. To choose a simulation period encompassing typical pollen release, we considered that most hemp cultivars are photosensitive, flowering as day lengths shorten below a threshold (10–12 h), following the summer solstice[36,37] which varies with latitude. An allergen study measured airborne *Cannabis* pollen counts for 5 years (1992–1996) in Omaha, Nebraska, finding pollen starting in the last two weeks of July, peaking in late August, and ending in mid-September[43]. A 2022 Colorado survey reported cross-pollination events occurring between July to mid-October[2]. Therefore, we chose to simulate dispersion from July into November, to see the continuation as weather conditions change.

Meteorological input

To drive the LS model, we use meteorological fields obtained from a Weather Research and Forecasting (WRF) model simulation over the CONUS for calendar year 2016[23]. This dataset comprises an hourly time series of meteorological conditions on a 12 km-square horizontal grid, and has been evaluated extensively in previous studies[24]. At the grid-point nearest to the centroid of each county, we extracted meteorological parameters describing horizontal wind shear, convection, boundary layer height, and surface roughness, namely, the five variables mentioned above, (u^*, L, w^*, z_0, z_i). We averaged these parameters across local noon and midnight hours for each month from July to November to form county-specific monthly average "day" and "night" cases.

Model simulations and boundary conditions

In each LS simulation—a daytime and a nighttime simulation for each county and for each month—100,000 particles were released at a height of 2 m with initial velocity selected from the velocity PDF, minus a constant settling velocity. Particles were removed from the simulation when they travelled above the boundary layer height z_i, upwind 10 m, or downwind 50 km. Pollen traveling above the ABL were considered to be subject to transport far beyond the 50 km bounds of our model domain. Such long-distance transport was not considered, as this study is more focused on exploring risk of cross-pollination from nearby farms. The downwind extent of the domain was determined by computational constraints (resolution of depositions of 100,000 particles, and simulation time for this number of particles to traverse the domain), while considering cross-pollination distances of interest (5 km, 10 km, 20 km and greater). Particles were considered to have "deposited" at a height of 1 m and were removed from the simulation. This height was greater than the surface roughness length for the majority of counties, the lowest permissible bound for the model which allows for comparison between counties. In summary, particles are released at a 2 m height, advected by the wind model, and are considered deposited when they fall below 1 m. Each simulation yielded a dispersal kernel, or (normalized) number of particles deposited downwind from the source, in 250 m wide bins.

Simplifications

To facilitate a large-scale comparative model, the simulation conditions are simplified. We treated dispersion for every county as if pollen was travelling over a flat, rough plane. The following phenomena and conditions are not considered: canopy escape, deposition probability, precipitation, topology, ground-cover, or variable source. We chose these simplifications to compare the effects of weather conditions on model predictions of dispersion across counties and seasons. We are primarily interested in how the spatio-temporal distribution in the five meteorological input parameters, described above, yield geographic and seasonal patterns in pollen transport distances. To get a nationwide overview, we chose to vary only these five parameters. For a more accurate assessment of local dispersion from an individual field, the other phenomena and conditions listed above need to be taken into account.

Data and code availability

Simulation results, monthly-averaged meteorological input data, and all dispersal model code are made available in the Virginia Tech Data repository: https://doi.org/10.7294/25718625.v1.

Received: 3 May 2024; Accepted: 20 August 2024
Published online: 04 September 2024

References

1. Malone, T. & Gomez, K. Hemp in the United States: A case study of regulatory path dependence. *Appl. Econ. Perspect. Policy* **41**, 199–214 (2019).
2. Colorado Department of Revenue's Marijuana Enforcement Division and Colorado Department of Agriculture. HB21-1301 Cross Pollination Working Group Technical Report. Tech. Rep. (2022). https://sbg.colorado.gov/med/1301-Work-Groups.
3. McCarty, T. & Young, J. Hemp production network effects: Are producers tipped toward suboptimal varietal selection by their neighbors?. *J. Appl. Farm Econ.* **3**, 4 (2020).
4. Habekost, Z. Defining Damage: The "Damage to Commercial Agricultural Products" Exception to Oregon's Right-to-Farm Law. *Or. L. Rev.* **100**, 181 (2021).
5. Small, E. & Antle, T. A preliminary study of pollen dispersal in Cannabis sativa in relation to wind direction. *J. Ind. Hemp* **8**, 37–50 (2003).
6. Faegri, K., Kaland, P. E., Krzywinski, K. *et al.Textbook of Pollen Analysis.* Ed. 4 (John Wiley & Sons Ltd., 1989).
7. Canadian Seed Growers' Association. Canadian regulations and procedures for pedigreed seed crop production. https://seedgrowers.ca/seed-growers/requirements/ (2023).
8. California Crop Improvement Association. Hemp Crop Standards (*Cannabis sativa*). https://ccia.ucdavis.edu/standards/crop-standards/hemp-crop-standards (2019).
9. Oregon CBD. Feminized Seed and the Ethics of *Cannabis* Farming. https://oregoncbdseeds.com/farmerresources/ (2017).

10. Cabezudo, B. *et al.* Atmospheric transportation of marihuana pollen from North Africa to the southwest of Europe. *Atmos. Environ.* **31**, 3323–3328 (1997).
11. Aznar, F. *et al.* Cannabis, an emerging aeroallergen in southeastern Spain (Region of Murcia). *Sci. Total Environ.* **833**, 155156 (2022).
12. Kuparinen, A., Katul, G., Nathan, R. & Schurr, F. M. Increases in air temperature can promote wind-driven dispersal and spread of plants. *Proc. R. Soc. B Biol. Sci.* **276**, 3081–3087. https://doi.org/10.1098/rspb.2009.0693 (2009).
13. Tackenberg, O., Poschlod, P. & Kahmen, S. Dandelion seed dispersal: The horizontal wind speed does not matter for long-distance dispersal - it is updraft!. *Plant Biol.* **5**, 451–454. https://doi.org/10.1055/s-2003-44789 (2003).
14. Soons, M. B., Heil, G. W., Nathan, R. & Katul, G. G. Determinants of long-distance seed dispersal by wind in grasslands. *Ecology* **85**, 3056–3068. https://doi.org/10.1890/03-0522 (2004).
15. Oneto, D. L., Golan, J., Mazzino, A., Pringle, A. & Seminara, A. Timing of fungal spore release dictates survival during atmospheric transport. *Proc. Natl. Acad. Sci.* **117**, 5134–5143. https://doi.org/10.1073/pnas.1913752117 (2020).
16. Savage, D., Barbetti, M. J., MacLeod, W. J., Salam, M. U. & Renton, M. Seasonal and diurnal patterns of spore release can significantly affect the proportion of spores expected to undergo long-distance dispersal. *Microb. Ecol.* **63**, 578–585 (2012).
17. Cimorelli, A. J. *et al.* AERMOD: A dispersion model for industrial source applications. Part I: General model formulation and boundary layer characterization. *J. Appl. Meteorol. Climatol.* **44**, 682–693 (2005).
18. Kuparinen, A. Mechanistic models for wind dispersal. *Trends Plant Sci.* **11**, 296–301 (2006).
19. Luhar, A. K., Hibberd, M. F. & Hurley, P. J. Comparison of closure schemes used to specify the velocity PDF in Lagrangian stochastic dispersion models for convective conditions. *Atmos. Environ.* **30**, 1407–1418 (1996).
20. Boehm, M. T. & Aylor, D. E. Lagrangian stochastic modeling of heavy particle transport in the convective boundary layer. *Atmos. Environ.* **39**, 4841–4850 (2005).
21. Boehm, M. T., Aylor, D. E. & Shields, E. J. Maize pollen dispersal under convective conditions. *J. Appl. Meteorol. Climatol.* **47**, 291–307 (2008).
22. Aylor, D. E. & Flesch, T. K. Estimating spore release rates using a Lagrangian stochastic simulation model. *J. Appl. Meteorol. Climatol.* **40**, 1196–1208 (2001).
23. US EPA. CMAQ Model Version 5.3 Input Data–1/1/2016- 12/31/2016 12 km CONUS, https://doi.org/10.15139/S3/MHNUNE (2019).
24. US EPA. Meteorological model performance for annual 2016 simulation WRF v3.8 (2019).
25. Karl, T. & Koss, W. J. Regional and national monthly, seasonal, and annual temperature weighted by area, 1895–1983 (1984).
26. Adams-Groom, B., Skjøth, C. A., Baker, M. & Welch, T. E. Modelled and observed surface soil pollen deposition distance curves for isolated trees of Carpinus betulus, Cedrus atlantica, Juglans nigra and Platanus acerifolia. *Aerobiologia* **33**, 407–416 (2017).
27. Sofiev, M. & Bergmann, K.-C. *Allergenic pollen: A review of the production, release, distribution and health impacts* (Springer, 2012).
28. Mendez, A. & Farazmand, M. Quantifying rare events in spotting: How far do wildfires spread?. *Fire Saf. J.* **132**, 103630. https://doi.org/10.1016/j.firesaf.2022.103630 (2022).
29. Aylor, D. E., Boehm, M. T. & Shields, E. J. Quantifying aerial concentrations of maize pollen in the atmospheric surface layer using remote-piloted airplanes and Lagrangian stochastic modeling. *J. Appl. Meteorol. Climatol.* **45**, 1003–1015 (2006).
30. Willis, G. & Deardorff, J. A laboratory model of diffusion into the convective planetary boundary layer. *Q. J. R. Meteorol. Soc.* **102**, 427–445 (1976).
31. Briggs, G. Plume dispersion in the convective boundary layer. Part II. Analyses of CONDORS field experiment data. *J. Appl. Meteorol. Climatol.* **32**, 1388–1425 (1993).
32. Stull, R. B. *An Introduction to Boundary Layer Meteorology* (Springer, 1988).
33. Choudhary, N., Siddiqui, M., Bi, S. & Khatoon, S. Effect of seasonality and time after anthesis on the viability and longevity of Cannabis sativa pollen. *Palynology* **38**, 235–241 (2014).
34. Aboulaich, N. *et al.* Variations and origin of the atmospheric pollen of Cannabis detected in the province of Tetouan (NW Morocco): 2008–2010. *Sci. Total Environ.* **443**, 413–419 (2013).
35. Farm Service Agency. FSA Crop Acreage Data Reported to FSA, 2023 acreage data as of August 9, 2023 (2023). https://www.fsa.usda.gov/news-room/efoia/electronic-reading-room/frequently-requested-information/crop-acreage-data/index.
36. Clarke, R. C. *Botany of the Genus Cannabis* (Haworth Press, Binghamton, NY, 1999).
37. Borthwick, H. & Scully, N. Photoperiodic responses of hemp. *Bot. Gaz.* **116**, 14–29 (1954).
38. Quaicoe, O., Asiseh, F. & Isikhuemhen, O. S. Qualitative analysis of industrial hemp production, markets, and sustainability in North Carolina. *United States. Agriculture* **13**, 887 (2023).
39. Coolong, T., Cassity-Duffey, K. & Joy, N. Role of planting date on yield and cannabinoid content of day-neutral and photoperiod-sensitive hemp in Georgia, USA. *HortTechnology* **33**, 138–145 (2023).
40. Malabadi, R. B. *et al.* Cannabis sativa: Botany, cross pollination and plant breeding problems. *Int. J. Res. Innov. Appl. Sci.* **8**, 174–190 (2023).
41. Steiner, J. Making hemp a 21st century commodity in Oregon-and beyond. *Crops Soils* **54**, 24–31 (2021).
42. Nieuwstadt, F. T. The turbulent structure of the stable, nocturnal boundary layer. *J. Atmos. Sci.* **41**, 2202–2216 (1984).
43. Stokes, J. R., Hartel, R., Ford, L. B. & Casale, T. B. Cannabis (hemp) positive skin tests and respiratory symptoms. *Ann. Allergy Asthma Immunol.* **85**, 238–240 (2000).
44. Millwood, R. *et al.* Pollen-mediated gene flow from transgenic to non-transgenic switchgrass (*Panicum virgatum L.*) in the field. *BMC Biotechnol.* **17**, 1–10 (2017).
45. Ecker, G., Meyer, T. & Auer, C. Pollen longevity and dispersion models for switchgrass. *Crop Sci.* **53**, 1120–1127. https://doi.org/10.2135/cropsci2012.06.0382 (2013).
46. Young, J. S. & McCarty, T. J. Adapting network theory for spatial network externalities in agriculture: A case study on hemp cross-pollination. *Am. J. Agr. Econ.* **105**, 1267–1287 (2023).
47. Aylor, D. *Aerial Dispersal of Pollen and Spores* (American Phytopathological Society, 2017).
48. Luhar, A. K. & Britter, R. E. A random walk model for dispersion in inhomogeneous turbulence in a convective boundary layer. *Atmos. Environ.* **23**, 1911–1924 (1989).
49. Luhar, A. K. The influence of vertical wind direction shear on dispersion in the convective boundary layer, and its incorporation in coastal fumigation models. *Bound.-Layer Meteorol.* **102**, 1–38. https://doi.org/10.1023/A:1012710118900 (2002).
50. Rodean, H. C. *Stochastic Lagrangian Models of Turbulent Diffusion* (Springer, 1996).
51. Beljaars, A. & Holtslag, A. Flux parameterization over land surfaces for atmospheric models. *J. Appl. Meteorol. Climatol.* **30**, 327–341 (1991).
52. Kantha, L. H. & Clayson, C. A. *Small Scale Processes in Geophysical Fluid Flows* (Elsevier, 2000).
53. Small, E. Evolution and classification of Cannabis sativa (marijuana, hemp) in relation to human utilization. *Bot. Rev.* **81**, 189–294 (2015).
54. Amarasinghe, P. *et al.* The morphological and anatomical variability of the stems of an industrial hemp collection and the properties of its fibres. *Heliyon* **8** (2022).
55. Borrell, J. S. Rapid assessment protocol for pollen settling velocity: Implications for habitat fragmentation. *Biosci. Horizons* **5**, 1–9. https://doi.org/10.1093/biohorizons/hzs002 (2012).

Acknowledgements
We thank D.E. Aylor and A.K. Luhar for their communications and guidance as we implemented the dispersal models from their papers. We thank X. Huang for extracting the meteorological data for each county. This research was supported in part by the Biotechnology Risk Assessment program of the USDA NIFA, under grant number 2019-33522-29989. Open access publication of this article was partially supported by Virginia Tech's Open Access Subvention Fund (OASF).

Author contributions
M.N., S.R., and H.F. designed the study, interpreted the results, and contributed to the writing of the manuscript. M.N. implemented the models, conducted simulations, analyzed the data, and produced the figures. All authors reviewed the manuscript.

Competing interests
The authors declare no competing interests.

Additional information
Supplementary Information The online version contains supplementary material available at https://doi.org/10.1038/s41598-024-70633-x.

Correspondence and requests for materials should be addressed to M.N.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Supplementary materials for *Cannabis pollen dispersal across the United States*

Manu Nimmala[1,*], Shane D. Ross[2+], and Hosein Foroutan[3+]

[1]Virginia Tech, Engineering Science and Mechanics, Blacksburg, Virginia 24061, USA
[2]Virginia Tech, Aerospace and Ocean Engineering, Blacksburg, Virginia 24061, USA
[3]Virginia Tech, Civil and Environmental Engineering, Blacksburg, Virginia 24061, USA
[*]nimmala@vt.edu
[+]these authors contributed equally to this work

Supplementary Methods

S0.1 Model Formulation

The Lagrangian Stochastic (LS) model is an application of Brownian motion to turbulent diffusion, in which the trajectories of many particles through the air are modeled as random walks. Each step of a particle's path is influenced by both random and deterministic motions, guided by the statistics of the local wind field. By releasing thousands of particles and computing an ensemble average of their trajectories, we can determine the relative concentration at any point in the domain and the mean shape of the plume.

In this study, we implement two LS model formulations for the Eulerian velocity pdf: a convective boundary layer (CBL) model for unstable conditions ($L < 0$)[1,2] and a surface layer (SL) model for stable conditions[3] ($L > 0$), where L is the Monin-Obukhov length.

The position increments for particles in the x (downwind) and z (vertical) directions are as follows[3],

$$dx = (u' + \overline{U})dt, \tag{1}$$
$$dz = (w' - v_s)dt, \tag{2}$$

where u' and w' represent the fluctuating horizontal and vertical velocities, $\overline{U}$ is the mean horizontal wind velocity described further in Section S0.2.1, and v_s is a constant settling velocity for hemp computed using Stoke's law to be 0.027 m/s based on a typical hemp pollen diameter of 30 μm[4,5]. As hemp pollen is nearly spherical[5], Stoke's law provides a good approximation of settling velocity[6,7].

Particle velocity increments[3] in the x and z directions are computed using the Langevin equation,

$$du' = a_u dt + b_u \mathcal{N}(0, dt), \tag{3}$$
$$dw' = a_w dt + b_w \mathcal{N}(0, dt), \tag{4}$$

which describes the incremental changes in u' and w' fluctuating particle velocities. The Langevin coefficients, a_u, a_w and b_u, b_w, account for the deterministic and stochastic components of particle acceleration, respectively. The stochastic timestep is drawn from a normal distribution with a mean of 0 and variance dt.

The timestep, dt is computed as a fraction[1,3] of the lagrangian timescale τ:

$$dt = 0.02\tau, \tag{5}$$
$$\tau = 2\frac{\sigma_w^2}{C_0\varepsilon}, \tag{6}$$

where we chose the constant $C_0 = 3$[8,9], σ_w^2 is the vertical velocity variance and ε is the turbulent dissipation rate.

The two model formulations employed differ in their computation of the Langevin coefficients a_u and a_w, and particularly in how they solve for the Eulerian fluid velocity pdf $P_E(u_i', z)$ in the equations[10] below.

$$a_i = \frac{\phi_i}{P_E} + \frac{\frac{1}{2}C_0\varepsilon\frac{\partial P_E}{\partial u_i}}{P_E}, \tag{7}$$
$$b_i = \sqrt{C_0\varepsilon}. \tag{8}$$

For both formulations, b_u and b_w remain the same as above.

S0.1.1 The Convective Boundary Layer (CBL) model for unstable conditions ($L < 0$)

The CBL model was introduced by Luhar et al. (1989)[11], and computes $P_E(u_i', z)$ as the sum of two Gaussian pdfs to represent convective updrafts and downdrafts in the boundary layer, shown in (9),

$$\begin{aligned} P_E &= AP_A + BP_B, \\ P_A &= \frac{1}{\sqrt{2\pi}\sigma_A} \exp\left(\frac{-(w' - \overline{w}_A)^2}{2\sigma_A^2}\right), \\ P_B &= \frac{1}{\sqrt{2\pi}\sigma_B} \exp\left(\frac{-(w' + \overline{w}_B)^2}{2\sigma_B^2}\right). \end{aligned} \tag{9}$$

It is extended to 2 dimensions based on Luhar (2002)[8,9], which takes the horizontal and vertical velocity fluctuations to be independent. The Langevin coefficients then become,

$$a_w = \frac{\phi}{P_E} - \frac{\frac{1}{2}C_0\varepsilon Q}{P_E}, \tag{10}$$

$$a_u = \frac{-u' C_0 \varepsilon}{2\sigma_u^2}. \tag{11}$$

The ϕ term has been adapted to heavy particles in Boehm et al. (2005)[2]. The full closure method to find A, B, $\overline{w_A}$, $\overline{w_B}$, σ_A, σ_B is shown in Luhar et al. (1996)[1]. These are functions of the wind velocity profiles, which vary with height and are described in Section S0.2. Although the original CBL LS model[1,11] was a one-dimensional model intended for the well-mixed boundary-layer, Boehm et al. (2008)[8] incorporated wind statistics into this model which transition smoothly from the surface layer to the convective boundary layer above.

S0.1.2 Surface Layer LS Model for Stable Conditions ($L > 0$)

For stable conditions, the SL model incorporates a Gaussian pdf (eq. (12)) with jointly Gaussian velocity components (u, v, and w), as described in[3,10,12], giving a P_E,

$$P_E = \frac{1}{(2\pi)^{3/2} \det \tau_{ij}^{1/2}} \exp\left(-\frac{1}{2}(u_i - U_i)\tau_{ij}^{-1}(u_j - U_j)\right). \tag{12}$$

where τ_{ij} is the mean Reynold's stress tensor, and contains terms for the variances and covariances of the wind velocity.

This yields the Langevin coefficients used in Aylor & Flesch (2001)[3].

$$a_u = \frac{1}{A} b_u^2 (\overline{u'w'}w' - \sigma_w^2 u') + \frac{1}{2}\frac{\partial \overline{u'w'}}{\partial z} + \frac{1}{A}(\sigma_w^2 \frac{\partial \sigma_u^2}{\partial z} u'w' - \overline{u'w'}\frac{\partial \sigma_u^2}{\partial z} w'^2 - \overline{u'w'}\frac{\partial \overline{u'w'}}{\partial z} u'w' + \sigma_u \frac{\partial \overline{u'w'}}{\partial z} w'^2), \tag{13}$$

$$a_w = \frac{1}{A} b_w^2 (\overline{u'w'}u' - \sigma_u^2 w') + \frac{1}{2}\frac{\partial \sigma_w^2}{\partial z} + \frac{1}{A}(\sigma_w^2 \frac{\partial \overline{u'w'}}{\partial z} u'w' - \overline{u'w'}\frac{\partial \overline{u'w'}}{\partial z} w'^2 - \overline{u'w'}\frac{\partial \sigma_w^2}{\partial z} u'w' + \sigma_u^2 \frac{\partial \sigma_w^2}{\partial z} w'^2), \tag{14}$$

$$A = 2(\sigma_u^2 \sigma_w^2 - \overline{u'w'}^2). \tag{15}$$

S0.2 Wind statistics

To compute the Eulerian velocity pdf $P_E(u_i', z)$, we need to specify the wind statistics at every point in the domain, i.e., the mean, variances, covariances, and skewness. Assuming stationarity and horizontal homogeneity, the wind field statistics remain constant over time and vary only with height. Under this assumption, boundary layer scaling techniques such as Monin-Obukhov similarity theory, mixed layer, and surface layer scaling can be employed to generate vertical profiles of wind statistics. As a result, only 5 meteorological parameters are required to drive the LS simulation: the friction velocity u^*, the Monin-Obukhov length L, the convective velocity scale w^*, the surface roughness length z_0, and boundary layer height z_i.

S0.2.1 Horizontal wind velocity profile

To model the mean horizontal wind-velocity profile, $\overline{U}$, we use the logarithmic wind velocity profile from Monin-Obukhov similarity theory[13] with the stability correction function, ψ_M.

$$\overline{U} = \frac{u^*}{0.4}\left[\ln\left(\frac{z}{z_0}\right) + \psi_M\right]. \tag{16}$$

For stable conditions, we use the stability function as reported in Beljaars & Holtslag (1991)[14], where $a = 1$, $b = 2/3$, $c = 5$, and $d = 0.35$. In this paper, they compare the resulting velocity profiles with field measurements and find that this parameterization performs well throughout the boundary layer despite the fact that surface-layer scaling is used. Optis et al. (2016)[15] also compared various stable wind profiles, including the one presented below, and show that it performs well up to 200 meters above the surface.

$$\psi_M = a\frac{z}{L} + b\left(\frac{z}{L} - \frac{c}{d}\right)\exp\left(-d\frac{z}{L}\right) + \frac{bc}{d}. \tag{17}$$

For unstable conditions, we use the stability function given by Paulson (1970)[16]. This has previously been used for other unstable LS simulations[3,8], and is considered to approximate measurements well[14].

$$\psi_M = -2\ln\left(\frac{1+\alpha}{2}\right) - \ln\left(\frac{1+\alpha^2}{2}\right) + 2\tan^{-1}(\alpha) - \frac{\pi}{2}, \tag{18}$$

where,

$$\alpha = (1 - 15\frac{z-d}{L})^{1/4}. \tag{19}$$

S0.2.2 Horizontal wind velocity variance

In stable conditions, we use the following relationship from Kantha and Clayson for the horizontal velocity variance (2000)[17],

$$\sigma_u^2 = 4u^{*2}\left(1 - \frac{z}{z_i}\right)^{3/2}. \tag{20}$$

In unstable conditions, we use the following parameterization from Luhar et al. (2002)[9] for the horizontal wind velocity variance.

$$\sigma_u^2 = (0.6w^*)^2. \tag{21}$$

S0.2.3 Vertical wind velocity variance

In stable conditions, for the vertical wind velocity variance, we use a relationship from Kantha and Clayson[17],

$$\sigma_w^2 = 3u^{*2}\left(1 - \frac{z}{z_i}\right)^{3/2}. \tag{22}$$

In the HYSPLIT model, this parameterization is provided as one option for simulating velocity variances in stable conditions. Oneto et al. (2020) compared dispersal results using the Kantha and Clayson (2000) scheme with other parameterizations offered by HYSPLIT, and found that there was little sensitivity.

In unstable conditions, we apply the merged parameterization from Boehm et al. (2005)[8]. This combines surface-layer scaling with that of the convective boundary layer, so that the conditions ranging from very unstable to neutral can be accurately modeled.

$$\sigma_{w,CBL}^2 = 1.7w^{*2}(z/z_i)^{2/3}(1 - 0.9z/z_i)^{4/3}, \tag{23a}$$

$$\sigma_{w,neutral}^2 = u^{*2}(1.7 - z/z_i), \tag{23b}$$

$$\sigma_{w,merged}^2 = \frac{(1-\exp(z/L))w^{*3}\sigma_{w,CBL}^2 + 5\exp(z/L)u^{*3}\sigma_{w,neutral}^2}{(1-\exp(z/L))w^{*3} + 5\exp(z/L)u^{*3}}. \tag{23c}$$

(23d)

S0.2.4 Turbulence kinetic energy dissipation rate

In stable conditions, we use the profile suggested by Rodean (1996)[10] for the entire stable boundary layer,

$$\varepsilon = \frac{u^{*3}}{0.4 * z}\left(1 + 3.5\frac{z}{L}\right)\left(1 - 0.85 * \frac{z}{z_i}\right)^{3/2}. \tag{24}$$

Rodean (1996) discusses that this profile was formed by fitting to a second-order turbulence model[18], and has generally agreed with measurements and other simulations.

In unstable conditions, we apply the merged surface layer/convective boundary layer profile described by Boehm et al. (2008)[8] to LS modeling, and found previously using Large Eddy Simulations[19],

$$\varepsilon = 0.4\frac{w^{*3}}{z_i} + \frac{u^{*3}(1-z/z_i)}{0.4z(1-15*z/L)^{1/4}}. \tag{25}$$

S0.2.5 Lagrangian Timescale

In all stabilities, we compute the Lagrangian time scale using[1,3],

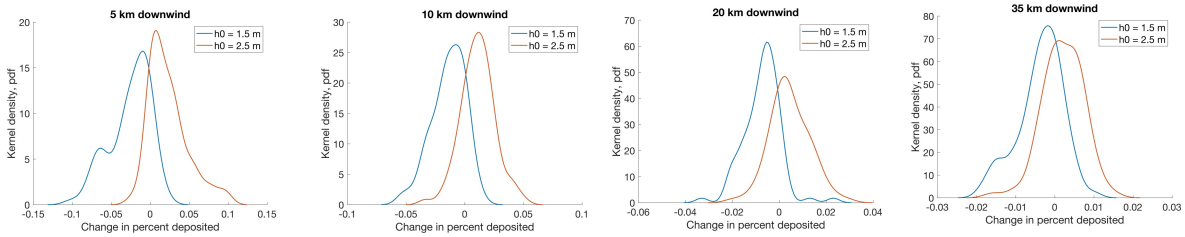
$$\tau = \frac{2\sigma_w^2}{C_0\varepsilon}. \tag{26}$$

Supplementary Figure S1

Sensitivity analysis by varying release height. In the main manuscript, we ran all simulations with a release height of 2 m. To explore sensitivity of our results to variations in release height, we randomly selected one county from each of the nine climate zones and reran simulations for day and night conditions for all months from July to November using release heights of 1.5, 2, and 2.5 meters. We found that reducing the release height resulted in less deposition throughout the domain for both day and night conditions, except in the first bin within 250 m from the source, where depositions increased. Increasing the release height had the opposite effect, with increased deposition throughout the domain. Figure S1(A) shows pdfs of the **change** in percent-deposited at 5, 10, 20, and 35 km downwind, while Figure S1(B) quantifies the change, showing mean and median differences and the percent-change at the same downwind distances.

While it is notable that changing the release height reduces depositions considerably, the shape of the deposition kernels after the first bin remains the same. This means that the spatial patterns we observe with 2 m release heights in the paper remain valid for this range of release heights. To demonstrate this, we also ran simulations for every county in the CONUS for day and night conditions, only for the month of July, using release heights of 1.5, 2, and 2.5 meters. In Figure S1(C), heat maps for each of these release heights show that although magnitudes change, spatial patterns remain the same. In our manuscript, we kept a fixed value of 2-meter release height in all simulations in order to focus on meteorological parameters. In our vulnerability analysis, we also incorporated land area and planted hemp acreage. However, to truly estimate cross-pollination risk and vulnerability, we recommend incorporating crop height, and additional factors such as location of farms, land topography, frequency of weather events like gusts or precipitation, and timing of pollen release.

(A)



(B)

Release height	5 km		10 km		20 km		35 km	
	Mean	Median	Mean	Median	Mean	Median	Mean	Median
1.5 m	-0.0256 (-21.306 %)	-0.0195 (- 21.196 %)	-0.0138 (- 24.576 %)	-0.0125 (- 24.272 %)	-0.0073 (- 28.411 %)	-0.0060 (- 25.000 %)	-0.0037 (- 31.787 %)	-0.0030 (- 27.273 %)
2 m	0	0	0	0	0	0	0	0
2.5 m	0.0229 (+ 15.696 %)	0.0170 (+ 13.281 %)	0.0110 (+ 15.627 %)	0.0105 (+ 15.328 %)	0.0044 (+ 13.383 %)	0.0035 (+ 10.294 %)	0.0020 (+ 13.416 %)	0.0020 (+ 12.500 %)

(C)

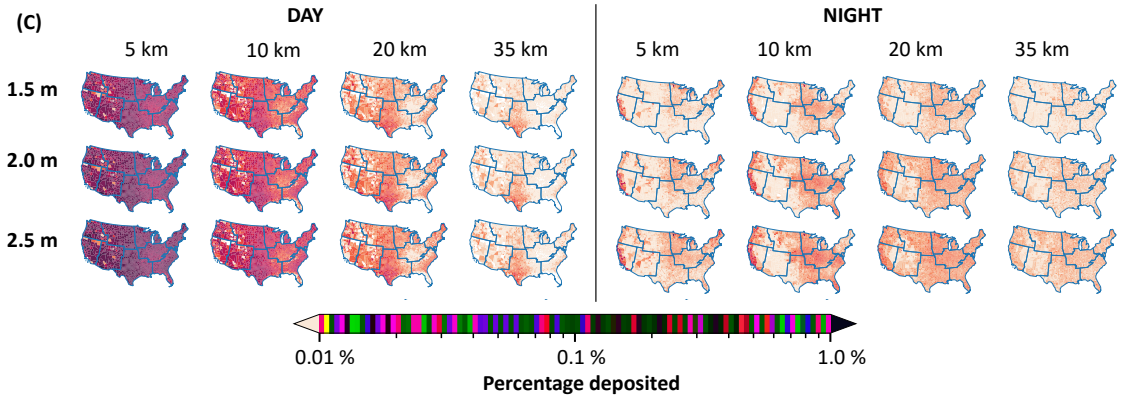


Figure S1. Sensitivity analysis conducted by varying the release height, where 2 m is the height originally used in the paper. (A) pdfs of the **change** (from a 2-m release height) in percent-deposited at 5, 10, 20, and 35 km downwind for both day and night conditions together. (B) Mean and median change in percent-deposited, and the percent-change for each release height, for both day and night conditions together. (C) Heatmaps showing depositions at 5, 10, 20, and 35 km distances from the source for three release heights for the month of July.

Supplementary Figure S2

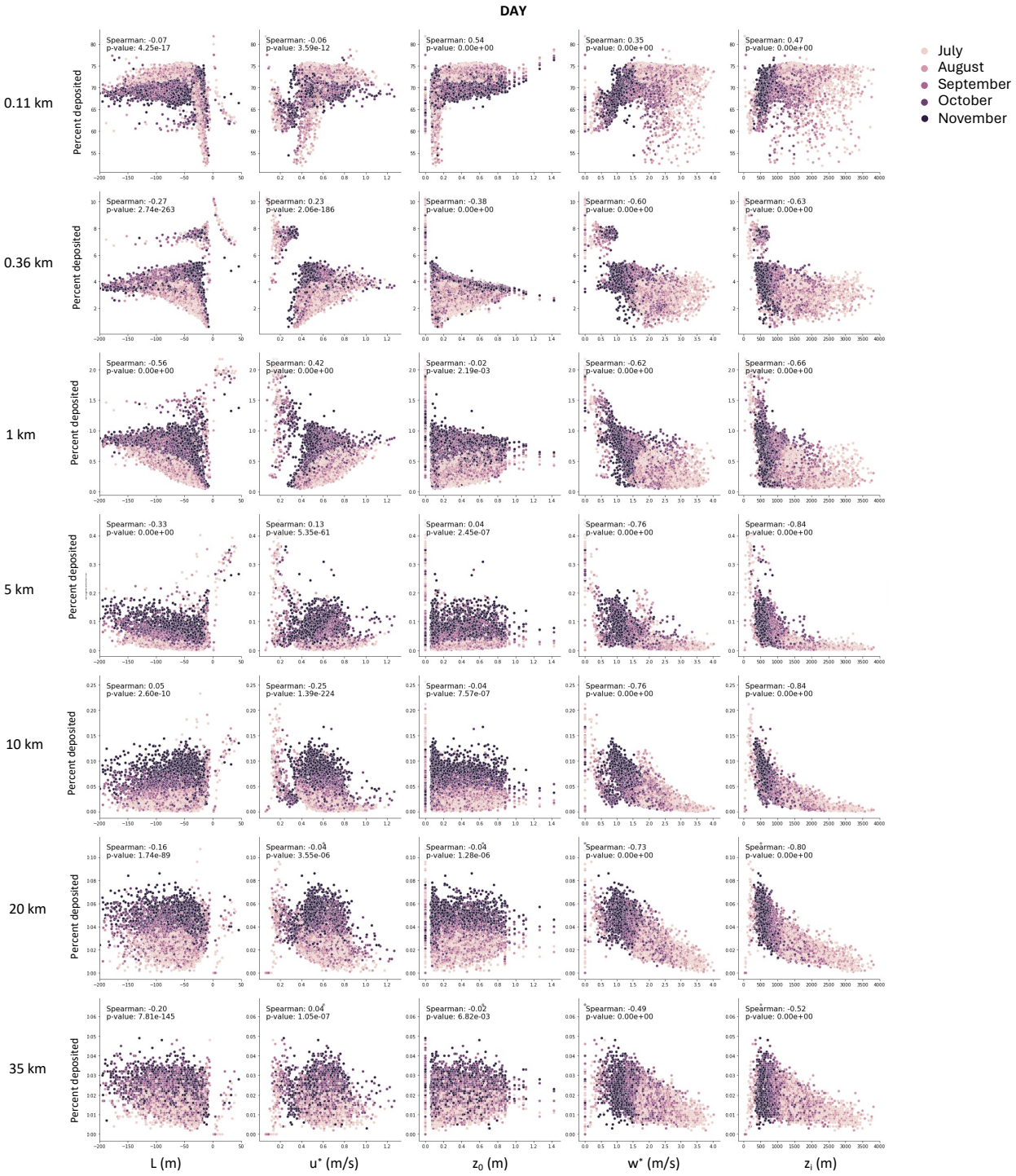


Figure S2. Scatterplots of five meteorological parameters for all day simulations vs. the percentage of particles deposited at distances downwind of the source. The Spearman correlation coefficients relating depositions at each downwind distance with the respective meteorological parameter are denoted for each plot. Decreasing deposition is most correlated with decreased boundary layer height z_i and w^* beyond 1 km from the source.

Supplementary Figure S3

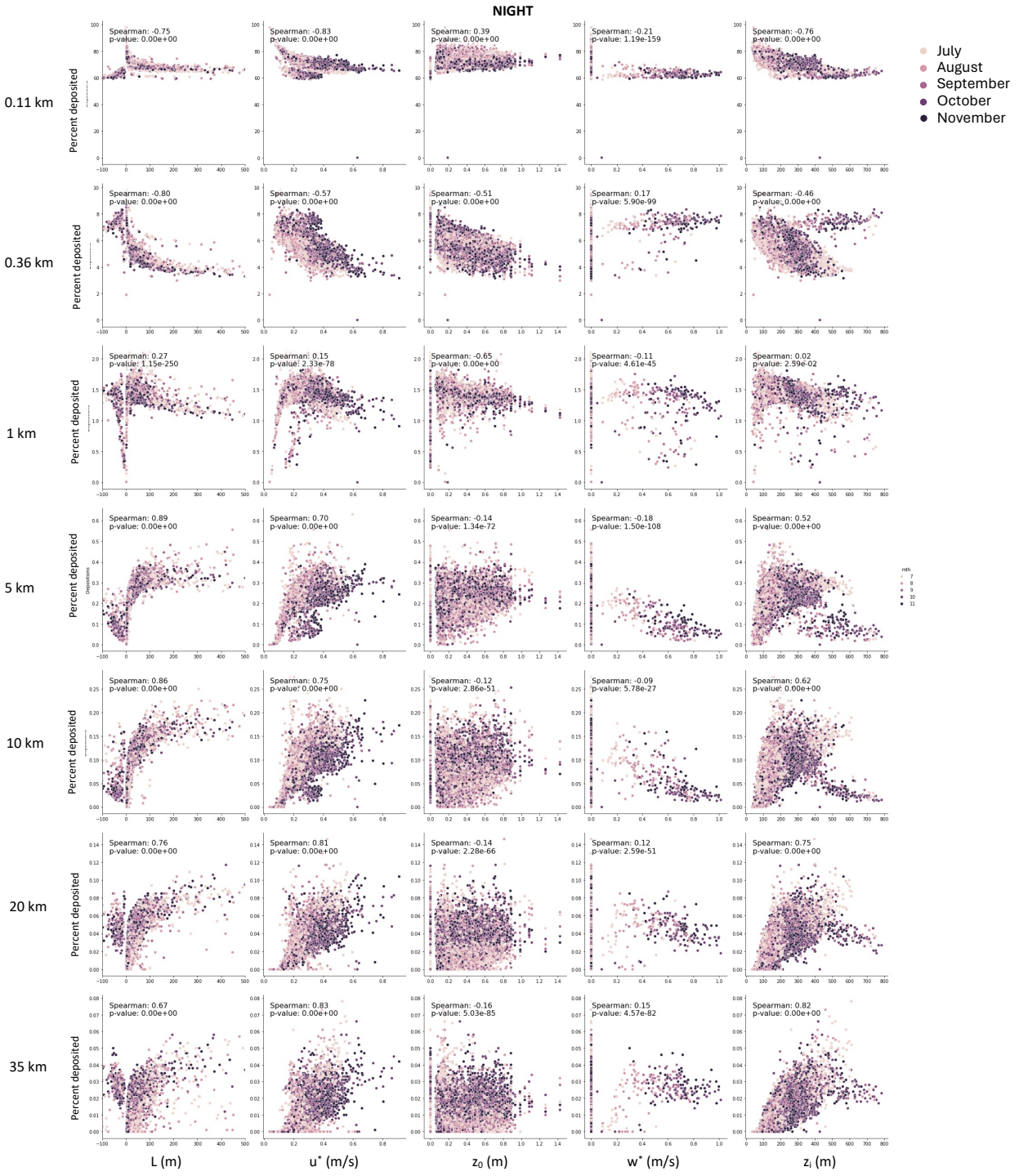


Figure S3. Scatter plots of meteorological input parameters vs. depositions for night cases. Scatterplots of five meteorological parameters for all **night** simulations vs. the percentage of particles deposited at distances downwind of the source. The Spearman correlation coefficients relating depositions at each downwind distance with the respective meteorological parameter are denoted for each plot. At night, greater boundary layer height z_i, friction velocity u^*, and obukhov length $|L|$ correlate with pollen travelling further - less deposition close to the source and increased deposition at all downwind distances beyond 1 km. The convective velocity scale, w^* is zero or a very small negative number for all night-time conditions, which make up the vast majority of nighttime case, and is not incorporated in the stable LS model.

Supplementary Figure S4

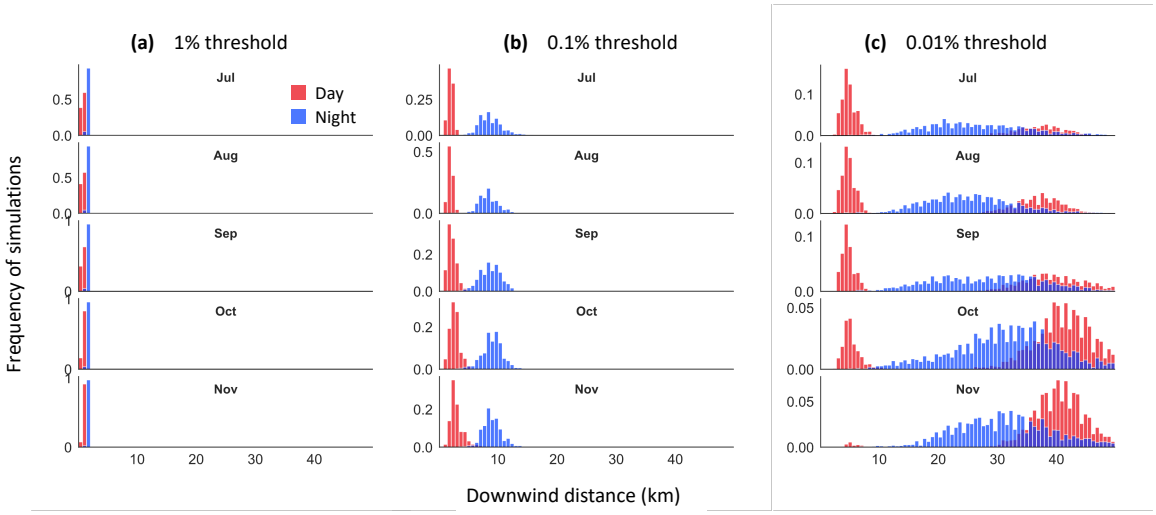


Figure S4. Distances at which dispersal kernels first fall below a threshold for each month: (a) 1%, (b) 0.1%, and (c) 0.01%. Red represents day simulations, while blue represents night. Seasonal variation is most pronounced for the 0.01% threshold distances, where the frequency of daytime distances beyond 30 km progressively increases from July to November.

Supplementary Figure S5

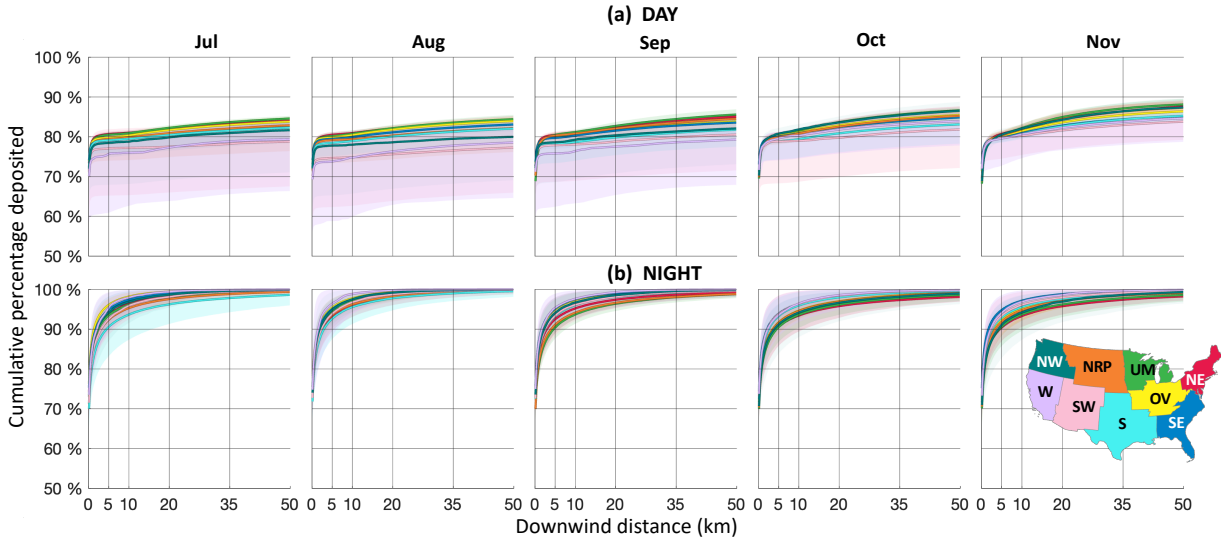


Figure S5. Median cumulative depositions for each month during (a) day and (b) night, separated by US climate region: Northeast (NE), Upper Midwest (UM), Ohio Valley (OV), Southeast (SE), Northern Rockies & Plains (NRP), South (S), Southwest (SW), Northwest (NW), and West (W). Shading represents data between the 10th and 90th percentiles. Note that the vertical axis is a log scale. There is a pronounced increase in total depositions in nighttime cases - most curves reach 100% within the domain. During the day, the kernels level out below 90%, although there is an increase in depositions from July to November.

Supplementary Figure S6

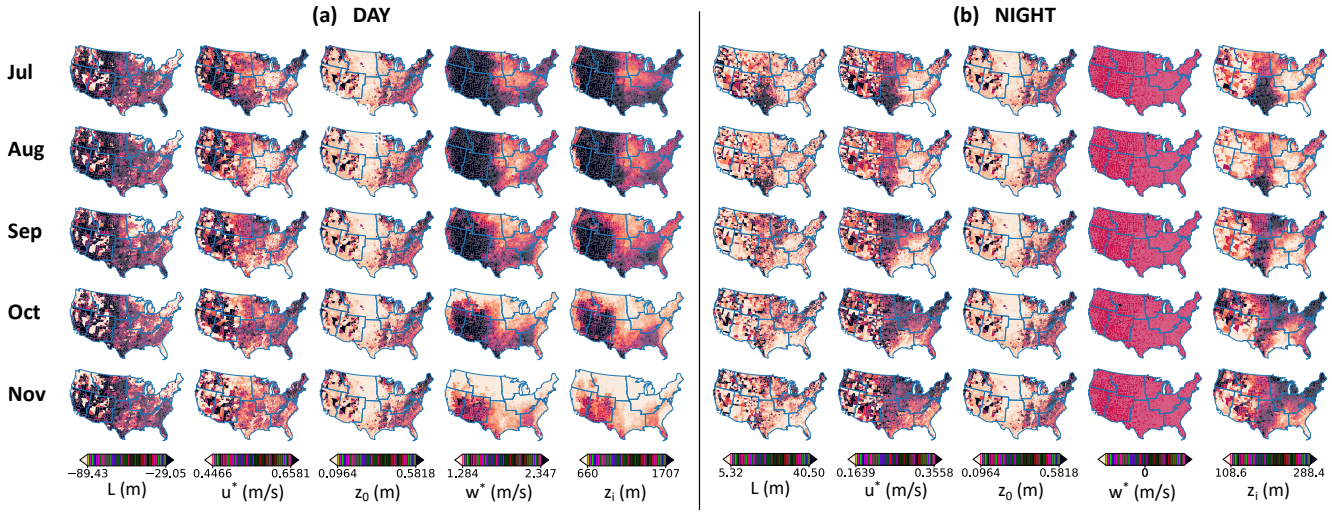


Figure S6. Heat maps of five meteorological parameters for all (a) daytime and (b) nighttime simulations over five months from July to November. The spatial and seasonal patterns visualized here mirror the deposition patterns shown in the main paper. During the day, the Southwest region maintains the highest convective velocity scale, w^* and boundary layer height, z_i throughout the season, and therefore the lowest daytime depositions overall. At night, the Southeast and Southwest regions have high friction velocity, u^*, high boundary layer height, z_i, lower roughness length, z_0, and high Monin-Obukhov length $|L|$, which results in less deposition in our simulations.

Supplementary Figure S7

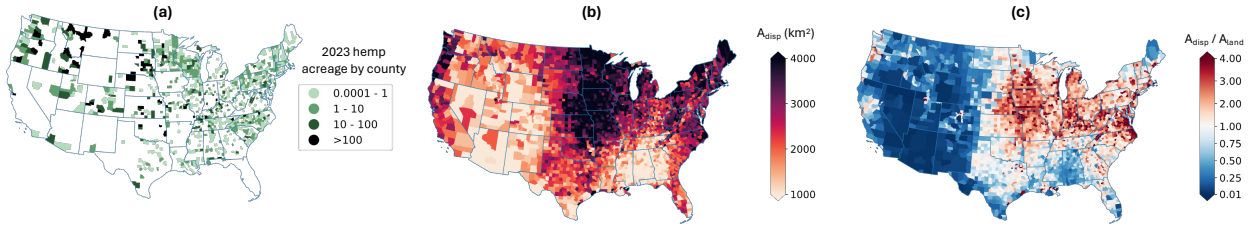


Figure S7. Components of the vulnerability metric. (a) The acreage of hemp, A_{hemp}, planted in each county as of 2023[20], where darker colors indicate greater planted hemp acreage. (b) The dispersal area A_{disp}, or area within a circle of radius equal to the average 0.01%-threshold distance, where darker colors indicate a greater dispersal area. (c) The ratio of A_{disp} to the land area of each county A_{land}, where red colors indicate regions where more $A_{disp} > A_{land}$.

Supplementary References

1. Luhar, A. K., Hibberd, M. F. & Hurley, P. J. Comparison of closure schemes used to specify the velocity pdf in lagrangian stochastic dispersion models for convective conditions. *Atmospheric Environ.* **30**, 1407–1418 (1996).
2. Boehm, M. T. & Aylor, D. E. Lagrangian stochastic modeling of heavy particle transport in the convective boundary layer. *Atmospheric Environ.* **39**, 4841–4850 (2005).
3. Aylor, D. E. & Flesch, T. K. Estimating spore release rates using a lagrangian stochastic simulation model. *J. Appl. Meteorol. Climatol.* **40**, 1196–1208 (2001).
4. Small, E. & Antle, T. A preliminary study of pollen dispersal in cannabis sativa in relation to wind direction. *J. Ind. Hemp* **8**, 37–50 (2003).
5. Clarke, R. C. *Botany of the Genus Cannabis* (Haworth Press, Binghamton, NY, 1999).

6. Borrell, J. S. Rapid assessment protocol for pollen settling velocity: Implications for habitat fragmentation. *Biosci. Horizons* **5**, 1–9, DOI: 10.1093/biohorizons/hzs002 (2012).

7. Aylor, D. *Aerial Dispersal of Pollen and Spores* (American Phytopathological Society, 2017).

8. Boehm, M. T., Aylor, D. E. & Shields, E. J. Maize pollen dispersal under convective conditions. *J. Appl. Meteorol. Climatol.* **47**, 291–307 (2008).

9. Luhar, A. K. The influence of vertical wind direction shear on dispersion in the convective boundary layer, and its incorporation in coastal fumigation models. *Boundary-Layer Meteorol.* **102**, 1–38, DOI: 10.1023/A:1012710118900 (2002).

10. Rodean, H. C. *Stochastic Lagrangian Models of Turbulent Diffusion*, vol. 6 (1996).

11. Luhar, A. K. & Britter, R. E. A random walk model for dispersion in inhomogeneous turbulence in a convective boundary layer. *Atmospheric Environ. (1967)* **23**, 1911–1924 (1989).

12. Thomson, D. J. Criteria for the selection of stochastic models of particle trajectories in turbulent flows. *J. Fluid Mech.* **180**, 529–556, DOI: 10.1017/S0022112087001940 (1987).

13. Stull, R. B. An introduction to boundary layer meteorology. *An introduction to boundary layer meteorology* (1988).

14. Beljaars, A. & Holtslag, A. Flux parameterization over land surfaces for atmospheric models. *J. Appl. Meteorol. Climatol.* **30**, 327–341 (1991).

15. Optis, M., Monahan, A. & Bosveld, F. C. Limitations and breakdown of monin–obukhov similarity theory for wind profile extrapolation under stable stratification. *Wind. Energy* **19**, 1053–1072 (2016).

16. Paulson, C. A. The mathematical representation of wind speed and temperature profiles in the unstable atmospheric surface layer. *J. Appl. Meteorol. Climatol.* **9**, 857–861 (1970).

17. Kantha, L. H. & Clayson, C. A. *Small scale processes in geophysical fluid flows* (Elsevier, 2000).

18. Brost, R. & Wyngaard, J. A model study of the stably stratified planetary boundary layer. *J. Atmospheric Sci.* **35**, 1427–1440 (1978).

19. Moeng, C.-H. & Sullivan, P. P. A comparison of shear-and buoyancy-driven planetary boundary layer flows. *J. Atmospheric Sci.* **51**, 999–1022 (1994).

20. Agency, F. S. FSA Crop Acreage Data Reported to FSA, 2023 acreage data as of August 9, 2023 (2023).