Bioaerosol Dispersal Across Scales: Regional Patterns, Field Study, and Model Evaluation

Manu Nimmala

Dissertation submitted to the Faculty of the
Virginia Polytechnic Institute and State University
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy
in
Engineering Science and Mechanics

Shane D. Ross, Chair
Hosein Foroutan, Co-chair
David D. Schmale
Anuj Karpatne

December 12, 2025
Blacksburg, Virginia

Keywords: Atmospheric pollen dispersal, Bioaerosols, Lagrangian Stochastic modeling

Bioaerosol Dispersal Across Scales: Regional Patterns, Field Study, and Model Evaluation

Manu Nimmala

(ABSTRACT)

Bioaerosols–including seeds, pollen, fungal spores, bacteria, and viruses–are fundamental agents connecting atmospheric processes to agriculture, ecosystem function, and human and animal health. This dissertation uses Lagrangian stochastic (LS) models to simulate how these particles travel and deposit across scales relevant for cross-pollination, with applications to many types of biological aerosols. First, we map seasonal and regional patterns of windborne hemp pollen across the United States by running LS models with weather data to simulate day- and night-time dispersal from summer through fall. These simulations identify areas more susceptible to cross-pollination and show how patterns shift across seasons and between day and night. We find regions more vulnerable to cross-pollination, with seasonal and diurnal shifting patterns in dispersal. Next, we work to detect and model genetically modified switchgrass pollen released from a small field in low-wind conditions during three sampling campaigns with a suite of novel samplers. We find that only our highest-volume samplers were able to detect pollen and that reducing the averaging window in the simulations substantially improved emission-rate estimates. Finally, we evaluate the 3D LS models used in this dissertation by comparing them to a high-fidelity model driven by large-eddy simulation (LES) in seven daytime convective boundary layer conditions. The LS models show moderate accuracy in strongly convective conditions, but they fail in near-neutral conditions due to issues in how they are parameterized rather than in their underlying equations.

Together, these results clarify when LS models can effectively substitute for more computationally intensive LES, reveal how sampler design and averaging choices shape what can be extracted from field measurements, and demonstrate the value of weather-aware modeling for cross-pollination risk assessment and broader questions of bioaerosol transport. Collectively, this work strengthens the scientific foundation needed to predict, manage, and mitigate the movement of biological aerosols in an increasingly variable atmosphere.

Bioaerosol Dispersal Across Scales: Regional Patterns, Field Study, and Model Evaluation

Manu Nimmala

(GENERAL AUDIENCE ABSTRACT)

Bioaerosols–including seeds, pollen, fungal spores, bacteria, and viruses–are fundamental agents connecting atmospheric processes to agriculture, ecosystem function, and human and animal health. This dissertation uses Lagrangian stochastic (LS) models to simulate how these particles travel and deposit across scales relevant for cross-pollination, with applications to many types of biological aerosols. First, we map seasonal and regional patterns of windborne hemp pollen across the United States by running LS models with large-scale weather data to simulate day- and night-time dispersal from summer through fall. These simulations identify areas more susceptible to cross-pollination and show how patterns shift across seasons and between day and night. Next, we work to detect and model genetically modified switchgrass pollen released from a small field in low-wind conditions during three sampling campaigns. We find that only the highest-volume samplers captured pollen, and that using shorter averaging windows in the simulations greatly improved emission-rate estimates. Finally, we evaluate the 3D LS models used in this dissertation by comparing them to a high-fidelity model driven by large-eddy simulations (LES) across seven daytime atmospheric conditions. The LS models show moderate accuracy in strongly convective conditions, but they fail in near-neutral conditions due to issues in how they are parameterized rather than in their underlying equations. Together, these results clarify when LS models can effectively substitute for more computationally intensive LES, reveal how sampler design and averaging choices shape what can be extracted from field measurements, and

demonstrate the value of weather-aware modeling for cross-pollination risk assessment and broader questions of bioaerosol transport. Collectively, this work strengthens the scientific foundation needed to predict, manage, and mitigate the movement of biological aerosols in an increasingly variable atmosphere.

Acknowledgments

I have been so fortunate. Over the last six and a half years, meeting by meeting, my two amazing co-advisors, Hosein Foroutan and Shane Ross, have shaped me into a more resilient person and a stronger researcher. This thesis is entirely a product of their time, patience, guidance, and encouragement. This work also would not have been possible without my collaborators who contributed to this research and were a delight to work with (including Hope Gruzweski, Regina Hanlon, Landon Bilyeu, Javier Gonzalez-Rocha, and David Schmale); the lovely and supportive Engineering Science and Mechanics department (Renee Cloyd, Mark Stremler); my labmates (including Albert Jarvis, Joshua Fitzgerald, Matt Werner, Kaustubh Gadamsetty, Mohamed Zakaria, Emad Masroor, Anjali Rawat, Xinyue Huang, Huan Yang, Carrie Carpenter, Nishan Pokhrel, Kaustuv Ray), cohort, and friends (Shilpa Thirukkovalur), who were always ready with both help and humor; and my supportive, enthusiastic family (my extremely wise aunts, my uncle who frequently gently turned my business ideas down and convinced me to finish, my cousins who kept me inspired with their fresh perspectives, and my grandparents with their excited expectations). Above all, I thank my mom, Uma Polisetti, who was always ready at any time with advice, support, love, help, and food. When I was sick, she flew across the country to take care of me. Her engineering career and curiosity for how things work inspired my own path. She is my hero. I am also grateful to my dad, Seshu Nimmala, for his guidance along the way. Finally, I thank my husband, Vanshaj Khattar, whom I met at the start of this journey and married by the end, and who has been my light through every tear and every triumph.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Bioaerosols are important and matter at different scales

Bioaerosols are small airborne particles formed and emitted from biological sources that link atmospheric transport to processes in climate, ecosystems, agriculture, and health. They include pollen, seeds, bacteria, fungi, and viruses, with diameters ranging from 1 nm to 100 μm [3]. They influence cloud microphysics and climate, plant reproduction and gene flow, and the spread of plant and human disease. For example, they are of vital interest to applications such as ice nucleation and cloud condensation [4], gene flow in relation to historical human cultivation [5, 6] and plant evolution [7], natural forest regeneration [8], airborne human infection [9], bioconfinement [10, 11, 12, 13] and regulation of genetically modified organisms (GMO) in agriculture [14, 15, 16], modeling ecological changes in plant populations [13, 17, 18, 19], allergenic pollen [20, 21, 22, 23], harmful algae blooms [24, 25, 26, 27], and the spread of plant diseases [28, 29, 30, 31, 32, 33]. In all of these settings, the central question is how these particles are emitted, transported, and ultimately removed from the atmosphere—that is, their dispersal.

The impacts of bioaerosol dispersal are scale-dependent, with relevant length scales ranging from a few meters to thousands of kilometers. In the context of human health, dispersal of bacteria and viruses via coughing and sneezing, such as those caused by chicken pox, measles, bacterial meningitis, and the famous 1.5 meter coronavirus social distancing rule

[34, 35] are all of interest within a few meters of the source. At several hundred meters to kilometers, legionnaire's disease, a potentially deadly respiratory illness caused by the spread of the aerosolized water containing Legionella bacteria, often spread from cooling towers becomes important [9, 35], as do diseases spread from animal farms tens of kilometers like avian influenza virus and Q-fever [9, 35]. Harmful algal blooms in oceans and lakes, such as the red tide in Florida [27] release toxins into the air upon aerosolization of water droplets, causing respiratory illness, and can be carried several kilometers inland [36].

For plant disease, Van der Heyden et al. [33] provides an extensive list of studies that monitor spores from infected crops from plot-scale experiments (e.g., apple scab, onion Botrytis, cucumber downy mildew) up to regional and national networks (wheat rust, Fusarium head blight, soybean rust). This monitoring is intended to prevent and manage local epidemics and large-scale spread, resulting in crop and financial losses [37]. Aylor [38] discusses the risk of spread of apple scab disease, from fungal spores released from decaying leaves on the orchard floor, escape from the floor, spread within the orchard, and spread between orchards [28, 38] causing deformed fruit, while tobacco blue mold has caused several regional-scale epidemics in the United States [28]. On the continental scale, Dillon and Dillon [35] discuss the introduction of soybean rust into North America from South America via Hurricane Ivan in 2004, with continued infestations following this event and resulting crop damage. Mohaimin et al. [37] argue that across these same local-to-continental scales, dispersal of crop-pathogenic bioaerosols is a major constraint on crop yield and food security [39].

Seeds are the larger than most other bioaerosols and far shorter dispersal distances. They tend to disperse within tens of meters and 100 meters is considered long-distance transport[40, 41, 42]. Yet, the study of their dispersal is essential for tracking gene flow for ecological applications like natural forest regeneration [8],the spread of plant and tree populations [19, 43, 44, 45, 46, 47, 48], and how climate change [49, 50] alters this spread.

Among bioaerosols, pollen has one of the most extensive dispersal literatures and is the primary focus of this dissertation. Pollen dispersal is studied primarily for gene flow prevention in agriculture and ecology. A major and current concern is the bioconfinement of genetically modified species, and prevention of their genetic drift through pollen dispersal at all scales [10, 11, 12, 13]. There has been extensive work on this for maize pollen [51, 52, 53, 54, 55] but other crops of interest include wheat and oilseed rape [56, 57]. Creeping bentgrass is an example of potential ecological change caused by gene flow from cultivated transgenic populations to feral populations [58, 59]. Similarly, switchgrass pollen dispersal is closely studied because it is often genetically modified and has the potential to alter ecosystems, particularly as it gains footing as a major biofuel source [11, 17, 18]. In our third chapter, we feature a field study which measures and models switchgrass pollen dispersal. In our second chapter, we study patterns in Cannabis pollen dispersal, specifically the long tails. Cannabis pollen is light weight and produced in large quantities [60], and so its dispersal and potential for cross-pollination at longer distances is of importance to the Cannabis industry. Long-range monitoring and trajectory analyses show that pollen can travel hundreds to thousands of kilometers, with birch and Cannabis pollen observed crossing national and continental boundaries [23, 61, 62]. Pollen dispersal is also tightly coupled to climate, with projections of changing allergen exposure or ecosystem shifts under future climate scenarios [20, 21, 22]. These studies illustrate that pollen dispersal spans the same local-to-regional scales as plant pathogens, and they motivate the multi-scale, Lagrangian modeling framework developed in this dissertation.

1.2 A framework for dispersal modeling

These scales, and the processes that dominate at each, motivate a range of modeling approaches for atmospheric dispersion. Most mechanistic atmospheric dispersion models can be viewed as two coupled components: (1) a description of the underlying flow and turbulence and (2) the method of simulating dispersion, a representation of how passive or heavy particles are transported within that flow.

1.2.1 Flow underlying dispersal

The development of the wind field for dispersal hinges on several key assumptions. Over what spatial and temporal scales is the dispersal occurring? At those scales, can stationarity be assumed, and if not, what time resolution is required? Can horizontal homogeneity of the wind field be assumed—is there terrain, is it flat, is it an idealized situation, or do other processes take over at larger length scales? What horizontal resolution is required to resolve those changes, the characteristics of the flow, and the characteristics of interest in the dispersal distribution? Can vertical homogeneity be assumed, or does surface wind shear need to be accounted for?

At mesoscales ($10^2 - 10^3$ km), dispersal models like HYSPLIT [63], CMAQ [64], and FLEXPART [65] are driven by numerical weather prediction models (e.g. WRF [66], MM5 [67], NCEP NAM [68]) and reanalysis products (e.g. ERA5 [69], MERRA-2 [70], NARR [71]) with horizontal grids on the order of 1 to 100 km [63].

For smaller scales (1 m $-$ 10^2 km), the diversity in wind-field development configurations explodes, each designed for a different combination of the key assumptions. At a high level, dispersal models at these scales can rely on measurements, modeling, or some combination

of the two. Wind fields via measurements can come from single sensors or multiple in the field providing points of data over time, profiles from a tower of sensors, a sodar, a weather balloon, or a drone, and a network of measurements (e.g. NOAA National Data Buoy Center (NDBC), METAR airport weather stations). CALMET for example, is a meteorological pre-processor that uses data such as this in combination with physics-based adjustments to produce a wind field for the dispersal model CALPUFF [72]. Modeling the wind field by solving the governing equations of fluid motion over a gridded domain (DNS, LES, RANS) can provide a much more detailed 3D wind field that could be time-varying, and can be customized to fit a micro-scale domain of a few meters, for example to resolve escape of particles from a plant canopy [73], or up to several kilometers, to model pollution over a city [74].

1.2.2 Dispersal

For the dispersion component, most mechanistic models represent particle motion in either an Eulerian or Lagrangian framework.

In a Eulerian framework, the transported quantity is the concentration $C(x, y, z, t)$ which evolves according to an advection-diffusion equation on a fixed grid. This approach is natural for coupling to chemistry, and underpins the large-scale chemistry-transport model CMAQ. At smaller scales, it struggles to resolve sharp gradients in concentration such as those occurring near the source [75], and would need a much finer grid in order to do so. However, Pan et al. (2014) reproduced observed particle concentrations inside the canopy roughness sublayer using this approach [73]. The familiar Gaussian plume model is a further simplification that assumes stationary horizontally homogeneous flow, a mean wind speed and eddy diffusivity that are the same everywhere in the domain, and a continuous point source, leading to an

analytical solution for Gaussian-shaped dispersion that can be shaped by a number of factors (the Pasquill-Gifford classes) [72]. The primary recommended dispersion model by the U.S. EPA, AERMOD, is a sophisticated extension of the basic Gaussian plume framework that incorporates vertical profiles of wind statistics, terrain effects, and convective turbulence, among others, into its prescriptions of Gaussian plume spread [76].

By contrast, the Lagrangian framework does not prescribe the shape of the plume in this way; it handles turbulence more naturally. In a Lagrangian framework, the model instead tracks the trajectories of individual particles through the wind field, and the ensemble average of their paths produces the concentration $C(x, y, z, t)$. Lagrangian methods handle dispersal near point and line sources and complex removal processes like wet and dry deposition quite easily, on a particle-by-particle basis [31, 75]. However, they can be difficult to scale up, requiring far more particles to resolve larger domains, dispersal in three-dimensions, and tail-end regions of low concentration. At regional to continental scales, widely used operational Lagrangian models include HYSPLIT [63] and FLEXPART [65], which transport large ensembles of particles on meteorological fields from numerical weather prediction or reanalysis products to simulate long-range dispersion and deposition [63, 77]. Similarly, CALPUFF's Guassian puff formulation, combines Lagrangian dispersal of a puff with a gaussian distribution to represent its expansion. [78]

Lagrangian stochastic (LS) models make up the bulk of modeling in the Lagrangian framework, in which particle turbulent velocities are modeled statistically. They are an application of Brownian motion to turbulent diffusion [79], in which each step of a particle's path is influenced by both random and deterministic motions, guided by the statistics of the local wind field. In pollen, seed, and spore dispersal modeling, they are most commonly assumed to be stationary within the dispersal time of an hour or less, horizontally homogeneous, and they output what is essentially a time-averaged plume. They take as inputs vertical profiles

of the fluctuating wind velocity statistics (variances and covariances) and mean horizontal velocities, and assume that the mean vertical wind velocity is zero. These are often constructed using boundary-layer scaling [54, 79, 80, 81, 82], given input parameters like 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. This framework has been applied extensively to pollen, spore and seed disperal [30, 40, 45, 46, 49, 83], and has also been extended to run in a time-varying capacity when driven by LES fields [84, 85].

Within this framework, two related but distinct LS formulations have been used most widely for daytime plant dispersal in the atmospheric boundary layer: a convective boundary-layer (CBL) model and a surface-layer (SL) model. The CBL LS model was originally developed for strongly convective conditions, in which turbulence in the bulk of the boundary layer is driven primarily by buoyancy. It represents the skewed fluctuating vertical-velocity distribution w' in the convective mixed layer with a bi-Gaussian probability density function [86]. As a result, it is able to reproduce plume-rise from the surface due to thermals and the subsequent descent from the boundary layer top as rising air cools [75, 85] . In this way, the CBL formulation can reproduce plume rise from surface sources and the subsequent descent of material lofted toward the top of the boundary layer. Although it has been modified to include some surface effects, it may under-predict concentrations near the source at ground-level [54]. Modifications have been made to incorporate surface layer effects into turbulence production for the CBL model. Luhar et al. [87] developed a closure scheme to model skewed turbulence that could reduce to Gaussian turbulence in the limit of zero skewness, closer to the form used for surface layer turbulence. [54] introduced a parameterization that merges surface layer wind statistics with convective boundary layer statistics. However, it does not include the covariance terms known to be important in the surface layer [54].

One other common formulation for LS models assumes a jointly Gaussian distribution for

the fluctuating velocity components (u', v', w'), which can then include the covariance $\overline{u'w'}$ between horizontal and vertical velocities [79, 88]. The assumption of the gaussian vertical velocities is well-suited for shear-dominated turbulence at the surface, and has produced good agreement with near-source measurements of pollen and spore dispersal [29, 31, 53, 75, 89]. In the chapters that follow, we refer to it as the SL (surface-layer) LS formulation, where it is mostly applied. While the SL model captures the spread and rise of the plume in unstable conditions, it cannot simulate the subsequent fall of the plume downwind. The CBL model would perform better at longer distances [54]. These tradeoffs motivate Chapter 4 of this dissertation.

1.3 Research Objectives

Aerobiology and atmospheric dispersion modeling have developed over decades into a vast, multidisciplinary field spanning microbiology, plant pathology, ecology, and atmospheric science—countless bioaerosol dispersal applications with impacts on a continuous spectrum of scales, simulated with an array of dispersal modeling techniques.

In light of this breadth of background, this dissertation focuses on three specific gaps in the literature: (1) quantification of cross-pollination risk for cannabis across seasons and geographical regions, (2) measurement and modeling of GMO pollen from a small source in low-wind conditions, and (3) validation of the two stationary LS models previously addressed against a high-resolution time-varying particle dispersal model driven with LES.

These works are tied together by the use of Lagrangian stochastic models in different spatial scales, driven by different meteorological inputs. Figure 1.1 illustrates the overall modeling framework for this dissertation.

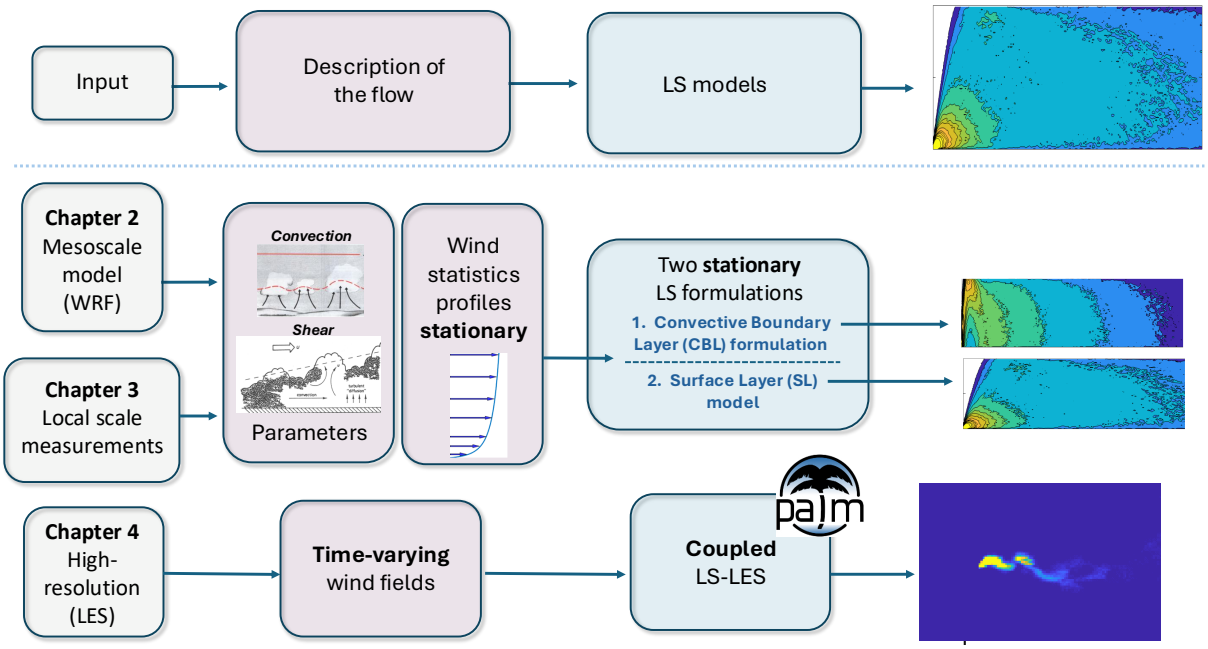


Figure 1.1: The LS modeling framework used in each chapter. In Chapters 2 and 3, we use mesoscale model inputs and local-scale meteorology inputs respectively to drive two stationary LS model formulations. In Chapter 4, we use the time-varying LS-LES simulations as a benchmark for the stationary LS formulations.

In Chapter 2, we use boundary-layer parameters and wind statistics from a mesoscale weather model to drive regional-scale, two-dimensional LS simulations on a 50 km grid, applied county by county across the continental United States. By coupling daytime convective-boundary-layer (CBL) and nighttime surface-layer (SL) formulations, this chapter quantifies how cannabis pollen dispersal potential varies across regions and seasons under realistic meteorological forcing. To the best of our knowledge, this is the first large-scale simulation study of the inhomogeneity of pollen dispersal across regions and seasons.

In Chapter 3, we focus on near-field dispersal from a very small, transgenic switchgrass field trial under predominantly low-wind conditions. Using a combination of novel measurement techniques, together with local-scale LS modeling (three-dimensional SL formulation, appropriate for local scales), this chapter examines pollen transport within roughly 25 m of the

source and uses it to estimate pollen source strength, identifying diurnal trends.

In Chapter 4, we compare the two stationary three-dimensional CBL and SL formulations driven with stationary wind statistics against PALM LES, a time-varying LS-LES simulation, in a range of daytime convective conditions to systematically evaluate them and identify when they are reliable and where they fail. To the best of our knowledge, there have been no studies which validate these models in three-dimensions in a range of convective regimes.

These three projects provide a cross-disciplinary and multi-scale view of bioaerosol transport. They combine tools from engineering mechanics (Lagrangian stochastic models and large-eddy simulation), atmospheric science (boundary-layer turbulence and regime transitions), and plant biology (hemp and switchgrass phenology, pollen traits, and cross-pollination outcomes). Throughout, the modeling is anchored by what can actually be measured in the field and by how dispersion information is used in practice by growers, regulators, and other stakeholders. The remainder of this dissertation returns to these themes, using the regional simulations, field campaigns, and LES comparisons together to ask not only how far and where pollen can travel, but also which levels of model complexity are appropriate for which types of questions.

Chapter 2

Cannabis pollen dispersal across the United States[1]

2.1 Introduction

The 2014 and 2018 US Farm Bills legalized the production of industrial hemp (Cannabis sativa) for cannabidiols, seed, and fiber[90]. 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 [91, 92]. Financial impacts reported in a 2022 Colorado survey[91] ranged from $12,000 to millions of dollars, with an Oregon lawsuit alleging damages of over $8 million [93]. Economic modeling[92] 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[90], 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 (2003) [60]. 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

[1]This chapter has been published as: Nimmala, M., Ross, S. D., & Foroutan, H. (2024), Scientific Reports 14:20605, doi:10.1038/s41598-024-70633-x.

excellent seed set", i.e., successfully cross-pollinate. The authors noted that due to its small size (∼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 [94]. 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 - 5 km [95, 96], but there have been reports of cross-pollination up to 20 km [97] and even 48-96 km away[91]. Two back-trajectory studies have demonstrated that Cannabis pollen likely travelled over 200 km, from Northern Africa to Spain [61, 62]. 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[49]. Another found that sustained updrafts caused dandelion seeds to disperse further, while horizontal wind speed did not play a factor [40]. In contrast, Soons et al (2004)[45] 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. [98] used the Hybrid Single-Particle Lagrangian Integrated Trajectory (HYSPLIT) 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. (2012)[30] 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. (2012)[30], 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 [76]), but these require more parameters and increase complexity [99].

We used two LS model formulations: a convective boundary layer model [54, 87, 100] for unstable (typically day) conditions and a surface layer model[31] 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 [101]. 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 [102]. 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.

2.2 Results & 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 2.1a and b show median day and night dispersal kernels on a log scale by month for each of nine US climate divisions [103], 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.

2.2.1 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, Figure

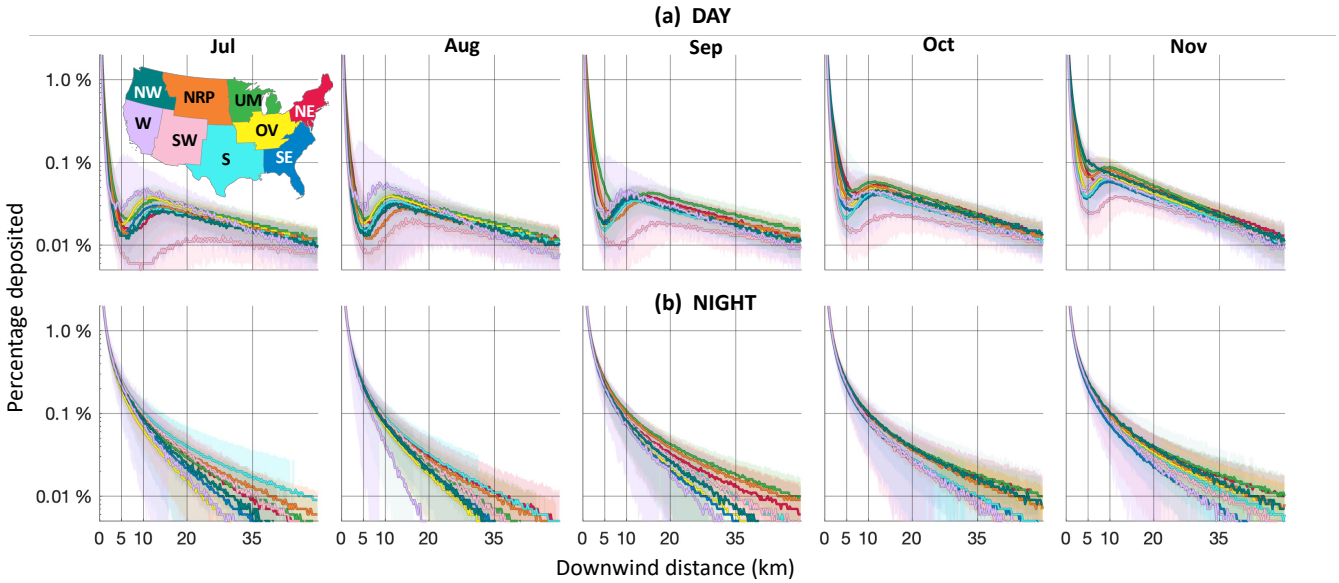


Figure 2.1: Median dispersal kernels for each month during (a) daytime and (b) nighttime, separated by US climate region: Northeast (NE), Upper Midwest (UM), Ohio Valley (OV), Southeast (SE), Northern Rockies & Plains (NRP), South (S), Southwest (SW), Northwest (NW), and West (W). Dispersal kernels are 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.

2.1 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 2.2a 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.

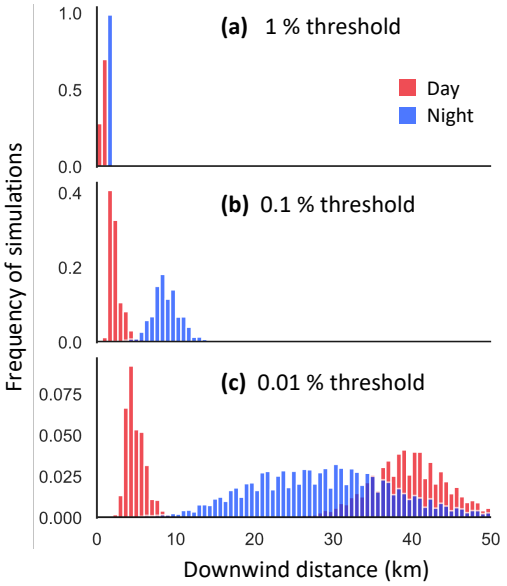


Figure 2.2: Distances at which dispersal kernels first fall below a threshold: (a) 1%, (b) 0.1%, and (c) 0.01%. Red represents day simulations, while blue represents night.

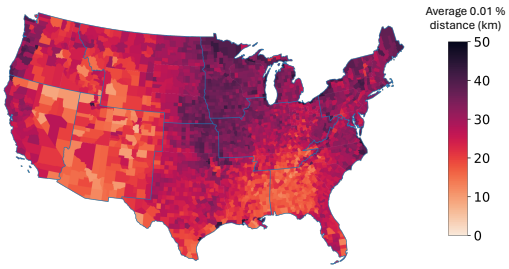


Figure 2.3: Heat map of 0.01%-distances averaged over all day and night simulations from July to November for each county.

While this steep decline in depositions appears to support commonly-used hemp isolation distances (< 5 km [95, 96]), even 1% of 100 million pollen grains would result in 1 million pollen depositing at that distance. In Figure 2.2b, 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 Figure 2.2c. This fat tailed deposition kernel is common for wind-dispersed

species[23, 104], and poses challenges when computing the risk of rare events in the tail, e.g., burning embers from a wildfire[105] or cross-pollination. For hemp in particular, the Small and Antle experiment [60] 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 [23, 53].

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 2.2b and c show stark differences between day and night simulations, driven by diurnal differences in wind conditions. For more detail, see Supplementary Figure S4. Below the 0.01% threshold, we 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.

2.2.2 Daytime seasonal and spatial patterns.

In Figure 2.1, 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, Figure 2.1 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 [106, 107], 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 Figure 2.4a 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 [108]. Scatter plots and correlation values between daytime depositions and these meteorological parameters are provided in Supplementary Figure S2 and the monthly heatmaps are shown in Supplementary Figure S6. High convective conditions in summer leads 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 makes long-distance transport of pollen more likely [30, 49, 98], 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 [30, 40, 49], particularly in combination with increased wind speed [45]. 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 0.45-0.65 m/s, or maximum variations in 10-m wind speed of 2 m/s. It is likely that averaging resulted in less

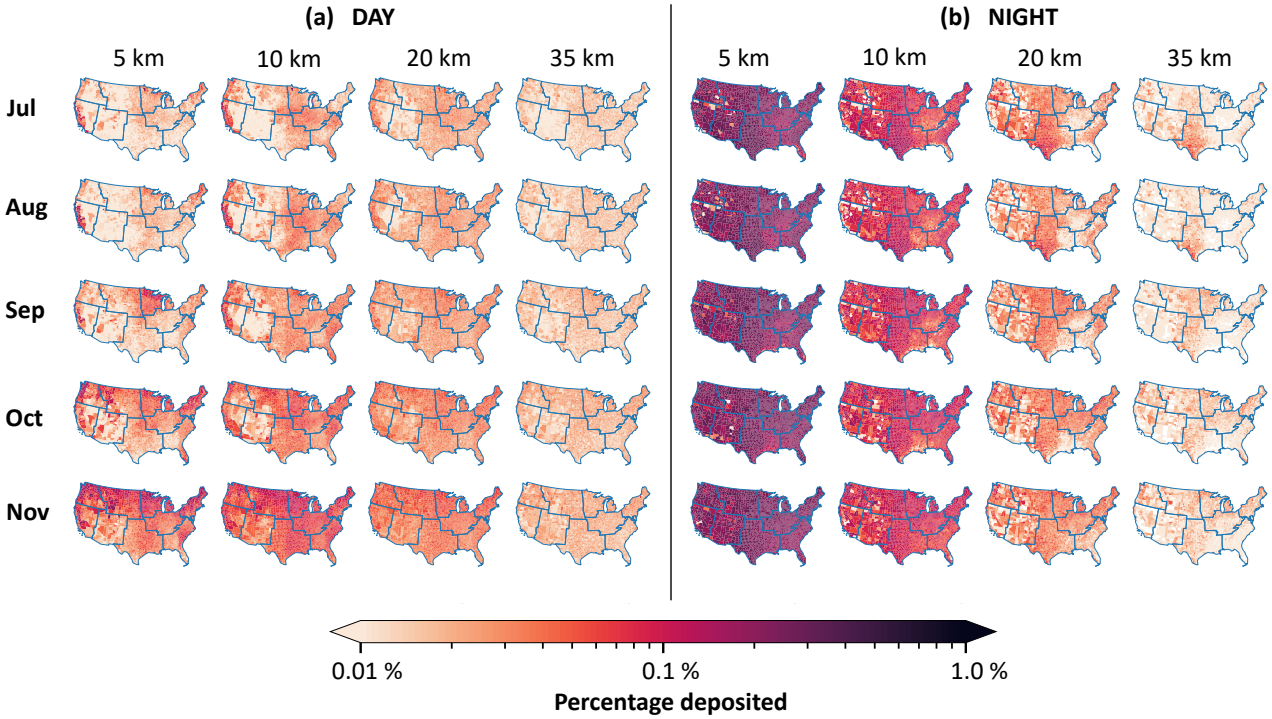


Figure 2.4: Percent of particles deposited in 250 meter-wide bins at downwind distances of 5, 10, 25, and 35 km for each county: (a) daytime simulations, (b) nighttime simulations. Note that the colorbar is a log scale.

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.

2.2.3 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. 2.1b. 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 2.4b 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 Figure S3. These parameters indicate more neutral conditions and greater wind shear, resulting in pollen travelling further from the source and depositing in greater amounts [45]. 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 [30] and remain airborne for longer [98] 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.

2.2.4 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 Figures 2.1 and 2.4. 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 Figure 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 Figure 2.1, where the daytime kernels have a shallower slope than and in Figure 2.4, 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. (2020)[98] 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 [98]. Choudhary et al. (2014) found that viability of Cannabis pollen only decreased substantially three days after release from the anther [109]. 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 hours to traverse the 50 km domain, and so 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 [23]. One study observed that male Cannabis anthers open and release pollen in the morning hours [109]. However, Cannabis pollen measurement studies found only slight diurnal changes in concentration [61, 110], 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.

2.2.5 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 Figure 2.3, 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 Figure 2.3. In Equation (2.1) 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 [2], 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 Figure 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,

$$\text{Vulnerability} = \frac{A_{\text{disp}}}{A_{\text{land}}} \times \frac{A_{\text{hemp}}}{A_{\text{land}}}. \tag{2.1}$$

Figure 2.5 shows a heat-map of the vulnerability metric for all counties with nonzero hemp acreage in 2023 [2]. 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 cross-pollination working group suggested a voluntary pinning system to track where hemp is planted in a region[91]. Rather than mandating specific isolation distances, we recommend a pinning system which includes location of outdoor planted, time of planting, and anticipated flowering dates. This could then 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.

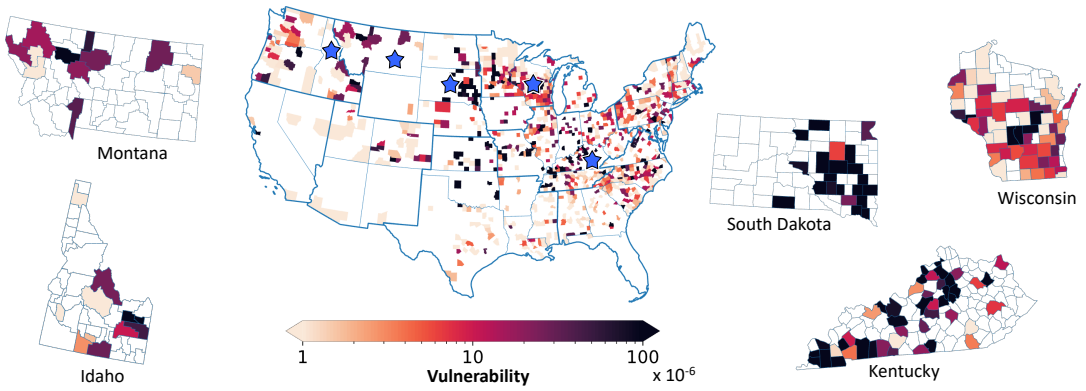


Figure 2.5: Vulnerability to hemp cross-pollination across the conterminous United States. The counties with non-zero planted hemp acreage as of 2023 are shown with darker shades showing greater vulnerability. The five states with the most land area with vulnerability greater than 10×10^{-6} are shown with stars.

The dependence of the vulnerability metric on dispersal distances and meteorological conditions tends to vary by region 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 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)[111, 112]. 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[113], and there are also non-photosensitive cultivars[114]. 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[110] 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[115].

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[116]. For example, in North Carolina, there was a crash in cannabidiol hemp production following a massive grower rush which exceeded demand[113]. 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[116]. These kinds of risks, in addition to the cross-pollination risk, can be managed with more intensive community planning.

2.3 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 [30, 98]. 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 [81, 108]. 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[117], in part because of its long flight times due to its small size compared to other pollen[60]. 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 [11], 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 [83]. 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[102], 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. [49] 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 [60] 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 [118]. 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.

2.4 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.

2.5 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 [31, 53, 75], 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 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. (1989 & 1996) [86, 87]. 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. (1996) [87] further intro-

duced 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. (2005)[100] adapted the model to include heavy particles, and Boehm et al. (2008)[54] 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 [87]. Predicted concentrations from the merged model were found to reasonably compare with measured aerial pollen concentrations [54].

Stable formulation. For all stable conditions, we used the surface-layer model as described in Aylor (2001)[31]. It differs from the CBL model in 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 [119]. 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 [108]. 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 [31]. The complete model formulations for both stable and unstable conditions can be found in Supplementary Methods 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 [54, 79, 80, 81, 82] 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 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 [111, 120]. A study examining hemp morphology found the mean height of 16 genotypes in the 1-2 m range [121]. We varied the release height by $\pm$ 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 can be found in the Supplementary Figure S1 available online. We used a settling velocity of $v_s = 0.027$ m/s, based on a typical hemp pollen diameter of 30 μm[60, 111], using Stokes' law. As hemp pollen is nearly spherical [111], Stokes' law provides a good approximation of settling velocity [75, 122]. Most hemp cultivars are photosensitive, flowering as day lengths shorten below a threshold (10-12 hours) following the summer solstice [111, 112], 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 [117]. A Colorado survey reported cross-pollination between July to mid-October [91]. 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 [101]. 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 [102]. 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.

Chapter 3

From Field to Sky: Measurement and Modeling of Transgenic Switchgrass Pollen Dispersal in the Atmosphere

3.1 Introduction

Accurate tracking and measurement of pollen dispersal in the atmosphere is important for assessing cross-pollination risks [89, 123], particularly in the case of genetically-engineered (GE) crops. Wind-dispersed pollen is the primary method of gene flow in many grasses, including switchgrass (Panicum virgatum), an important bioenergy crop [23]. It is a perennial, warm-season C4 bunchgrass found across most of eastern North America—from northern Mexico to southern Canada. Originally adopted as a forage crop, it is now a leading candidate for large-scale lignocellulosic biofuel feedstock in the U.S. and beyond [124]. There is increasing concern that the rapid growth and development of switchgrass as a biofuel could result in gene flow from GE switchgrass fields to nontransgenic fields (including wild populations), leading to both financial and ecological damage [10, 11, 17, 19, 125]. These changes could be compounded by the effects of climate change, where rising temperatures result in altered native switchgrass territory [18]. Therefore, there is an urgent need for field experiments and modeling efforts to characterize the dispersal of airborne switchgrass pollen in relation

to meteorological factors for regulation and risk management purposes.

There are limited experimental and modeling studies on switchgrass gene flow [19]; these model pollen dispersal with and without wind breaks [126], experimentally quantify the dispersal and cross-pollination of transgenic switchgrass [11], and model transport in low and high-wind conditions [83]. In 2011, Millwood and colleagues conducted the first regulated transgenic switchgrass field experiments in the U.S. [11]. A 3-year field experiment was performed in Oliver Springs, Tennessee, U.S.A. where 100 clonal switchgrass Alamo plants transgenic for an orange-fluorescent protein (OFP) were used as the pollen source (whole plants, including pollen, were orange-fluorescent). To assess pollen movement, pollen traps were placed at 10 m intervals from the pollen-source plot in the four cardinal directions extending up 100 m from the field. Results showed that pollen-mediated gene flow is likely to occur over distances of at least 100 m [11]. This study provided important baseline data useful to determine isolation distances and other management practices, should transgenic switchgrass be grown commercially in relevant environments. Since switchgrass is an obligate outcrossing perennial grass, there are concerns about gene flow and the need for bioconfinement, especially for pollen [10, 19, 127]. Moreover, since North America is the geographic center of switchgrass diversity, a better understanding of pollen movement in this species is needed [19].

The spread of pollen through the atmosphere involves processes of liberation, drift, and deposition [89, 128]. Knowledge of these processes can help growers and producers make informed management decisions regarding pollen transport in seed production fields and neighboring farms [128]. Although atmospheric transport models can predict pollen movement, they often fail to incorporate actual measurements of pollen concentrations and viability. Various unmanned aircraft systems (UASs or drones) have previously been used to detect and monitor pollen movement over long distances in the lower atmosphere. Gottwald and Ted-

ders pioneered the collection of pollen with UASs [129]. They modified a remote-controlled biplane platform with two rotating drum samplers to collect pollen and plant pathogen spores over pecan and peach orchards. Their study demonstrated the significant potential for regional-scale transport of pollen and plant pathogens among orchards. Two decades later, Aylor and colleagues [53] combined ground-based sampling devices with UASs to collect pollen within and above a cornfield. Over the past decade, Schmale and colleagues have integrated autonomous systems into UASs, enabling teams of vehicles to coordinate flight missions and perform complex atmospheric sampling tasks [130, 131].

The allergen-management community needs a fast and reliable sensor network to measure airborne pollen concentrations to enable timely and accurate allergen reporting [132, 133, 134, 135]. Current allergen information reports only broad species group concentrations, typically at a daily resolution at best [132, 135]. Future airborne pollen forecasts can be enhanced by integrating known pollen emissions with large-scale atmospheric models. Understanding diurnal pollen release patterns could aid in allergen treatment and improve emission source data for potential forecast models [132]. To our knowledge, most airborne pollen field studies and corresponding allergen reports rely on Hirst-type samplers [136]. These sampling devices are constrained by a relatively low sampling rate of approximately 10 L/min [136, 137], necessitating either high airborne pollen concentrations or extended sampling durations to accurately characterize pollen levels. The latter constraint contributes to the coarse temporal resolution of daily allergen pollen reporting.

We hypothesized that (1) wind-dispersed pollen from switchgrass could be tracked and quantified using orange fluorescent protein (OFP) expression, (2) a Lagrangian Stochastic (LS) dispersal model could estimate pollen source strength in the field, and (3) an array of novel samplers could serve as viable alternatives to standard Hirst-type samplers. To test these hypotheses, we conducted a series of unique release-recapture field studies using GE switch-

grass in Oliver Springs, TN. Two hundred plants from five transgenic lines of switchgrass (Panicum virgatum L. 'Performer') were planted at the center of a clear-cut field. One block consisted of 100 plants expressing OFP under the control of a maize ubiquitin promoter (PvUBI1), while the other block contained 100 plants expressing OFP driven by a maize pollen-specific promoter (Zm13). Pollen from the atmosphere surrounding these blocks of transgenic switchgrass was collected using a series of fixed (ground-based) and mobile (drone-based) sampling devices at various distances from the field center. The efficacy of these various samplers was evaluated within 25 m of the source and up to 1 km from the source. LS dispersal simulations were conducted for pollen sampling periods using high-resolution wind measurements collected near the field. Pollen emission rates were estimated by combining simulated concentrations with field concentration measurements. By integrating high-resolution measurements and simulations, our study evaluates the performance of emerging sampling technologies and highlights their implications for biosecurity, allergen tracking, and ecological modeling.

3.2 Methodology

3.2.1 Field Site and Pollen Source

Field site

To assess the transport of wind-dispersed pollen from transgenic switchgrass plants, a twoyear field study was conducted under USDA APHIS BRS release permits (21-094-103r and 124-86SS5F1). The experiments were carried out at the Tennessee Agricultural Experiment Station, near the University of Tennessee's Forest Resources Research and Education Center at the Cumberland Forest Unit in Oliver Springs, TN (36.0483147, -84.4811417).

The field site was selected to satisfy the primary requirements for regulatory for transgenic pollen dispersal experiments: isolation and traceable source attribution. It was situated on recently cleared forest land, with felled trees surrounding a rough glade area. The site provided sufficient open area (∼1.5 ha) for switchgrass cultivation and sampler deployment, while heavily forested borders served as a natural barrier that reduced the likelihood of cross-contamination with nearby wild or cultivated switchgrass and limited off-site pollen transport. The field location was intentionally chosen in a remote, concealed area beyond a secured access point, ensuring restricted visibility and access. The field plot measured 45 ft × 30 ft, enclosed within a protective 50 ft × 65 ft fenced perimeter, as shown in Figure 3.1A. The outer fence was locked to prevent animal intrusion.

Transgenic line generation, analysis, and selection

Transgenic switchgrass plants expressing OFP were created by genetically engineering embryogenic callus derived from switchgrass seeds obtained from Ernst Conservation Seeds, Inc. (Meadville, Pennsylvania, USA). This was achieved through Agrobacterium-mediated transformation (Agrobacterium tumefaciens strain EHA105) as detailed in [138], using one of two binary plasmid constructs.

The first plasmid, pANIC10A-OFP [139], contained the hygromycin phosphotransferase (hph) selectable marker gene under the control of the switchgrass ubiquitin 2 (PvUbi2) promoter as well as an orange fluorescent protein (OFP) gene pporRFP under the control of the switchgrass ubiquitin promoter (PvUBI1). This promoter drives the expression of OFP in leaves, stems, and pollen.

The second plasmid, PSYBIN1aZm13pporRFP, also contained the OFP gene pporRFP under the control of a maize pollen-specific promoter (Zm13). This promoter drives the expression

of OFP in the pollen. This construct also contained a second OFP gene mOrangeER under the control of the cauliflower mosaic virus (CaMV) 35S promoter which enable the expression of this OFP in callus and green tissues. In addition, the plasmid also contained the hygromycin phosphotransferase (hph) selectable marker gene under the control of the PvUBI2 promoter. Several transgenic OFP-expressing shoots were recovered from hygromycin selection media (100 mg/L), and once rooted, plants were grown an environmental-controlled growth room (16/8 h day/night and 24/22řC day/night).

To confirm the presence of the OFP gene in the transgenic plants, PCR-screening was performed using primers (forward primer: GCAAAGTGGGGTCAAAGATG; reverse primer: CACCTTCAAGCCCTTCTTTG) designed to amplify a 556 bp fragment of the pporRFP gene. PCR-confirmed transgenic plants were moved to a greenhouse and grown (16/8 h day/night and 28/22řC day/night) until flowering. To identify transgenic events expressing OFP in pollen, visual analysis of OFP fluorescence was conducted on pollen grains from each event using epifluorescent microscopy as described by [140]. Transgenic lines in which all pollen grains exhibited OFP expression were propagated in the greenhouse and subsequently used in field experiments.

Planting

The planted area, less than 0.1 ha, consisted of 20 rows with 20 switchgrass plants per row, totaling 400 transgenic switchgrass plants arranged in a randomized design. These plants were hand-transplanted in the field at 76.2 cm intervals on three different dates. On July 20, 2021, 100 pANIC10A-OFP switchgrass plants from five transgenic events (20 clonal replicates per event) were transplanted. On August 26, 2021, another 100 PSYBIN1aZm13pporRFP plants from five transgenic events (20 clonal replicates per event) were transplanted. Lastly, on June 20, 2022, an additional 200 pANIC10A-OFP switchgrass plants from ten transgenic

events (20 clonal replicates per event) were transplanted in the field site. These last 200 plants were not mature enough to produce pollen during the field experiments. Figure 3.1B illustrates the locations of these plants and their ages in weeks during the August 2-3, 2022 field experiment. This experimental design was structured to monitor and analyze the dispersal from transgenic pollen over time and distance.

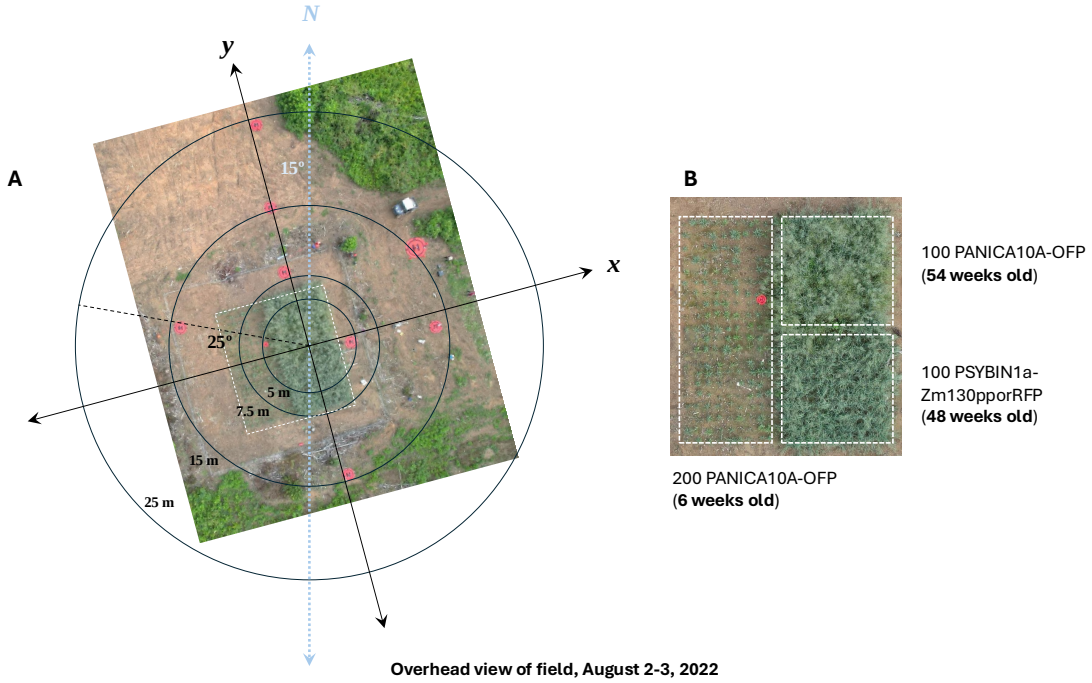


Figure 3.1: Top-down drone image of the field during field experiments conducted on August 2-3, 2022. (A) The field of GE switchgrass is outlined by a white dotted square and enclosed by a perimeter fence, as required by the APHIS BRS permits. Reddish-orange circular pads mark the locations of pollen sampling devices, positioned at increasing distances from the center of the source field. (B) A close-up view of the 45' × 30' field of GE Switchgrass, showing the locations of both strains of OFP-expressing switchgrass plants. Labels indicate the plant positions and ages during the August 2-3, 2022, field campaign.

3.2.2 Sampling Methods

Four types of volumetric particle samplers, shown in Figure 3.3, were used to capture GE switchgrass pollen and estimate the concentration of airborne pollen at various times and

distances from the source field. Each sampler had different sampling rates and sensing capabilities. To optimize equipment placement, wind forecasts and local conditions were assessed before the first sampling period. The samplers were strategically positioned based on prevailing and predicted wind directions, ensuring placement downwind of the expected pollen dispersal path. Figure 3.2 illustrates the placement of samplers around the field on each sampling day. The samplers were placed on reddish-orange circular drone landing pads to mark their locations and enhance visibility in overhead footage, as shown in Figure 3.1A for the August 2-3, 2022 field campaign.

Ground-based high-volume samplers (ED)

In anticipation of low pollen emission rates, several high-volume filter-based samplers (Science First #15000, Yulee, FL) [1] were deployed during the campaign (Figure 3.3A). Originally designed for educational use in schools, these samplers are referred to as "ED" samplers throughout the manuscript.

The barrel-shaped ED samplers draw air through a filter surface at an initial flow rate of 600 L/min [1]. Cellulose filters with a pore size of 11 μm and a diameter of 125 mm were used to collect airborne pollen and other atmospheric particles at 0.432 m above ground level. The ED samplers' volumetric sampling rate is 1000 times that of the IMP and DRN samplers, and 35 times that of the FRM sampler. This significant increase in sampling capacity allowed for improved detection of airborne pollen, particularly in cases of low pollen emission rates.

Ground-based medium-volume sampler (FRM)

A single near-Federal Reference Method (FRM) sampler (ARA Instruments, Eugene, Oregon) was deployed during the field experiments, shown in Figure 3.3B. This battery-operated

device samples air at a flow rate of 16.7 L/min. The unit is equipped with a filter sampler (PM_{10} filters were used in this study), meteorological sensors, and a particle counter. Additional details about this instrument are available on ARA's website [141].

Ground-based low-volume samplers (Impingers or IMP)

To evaluate the effectiveness of impinger-type samplers, three custom-designed impinger packages were deployed during the field campaign (Figure 3.3C) . These sampling packages are referred to as "IMP" throughout the manuscript.

The IMPs were constructed from high-density polyethylene, following the design specifications outlined in [24]. The 3D-printing (.stl) files for the impinger units are publicly available online [142]. These files were modified to accommodate a 15 mL polypropylene conical collection vial (Corning #CLS430791) and a stainless-steel tube with a 4 mm diameter opening.

The IMP samplers were mounted on a tripod 2 m above ground level to approximate the height of the switchgrass panicles, the open flower structures that produce pollen. The IMPs sampled airborne particles at a rate of 0.6 L/min, with collected particles entrained in sterile 15 mL conical tubes containing 2 mL of sterile deionized water.

Drone-based low-volume sampler (DRN)

To measure airborne pollen concentration at different altitudes above and downwind of the field, we used a drone-based sampling system consisting of the IMP unit mounted on a DJI Inspire 2 drone platform (Figure 3.3D). The system is described in detail in [25].

A key design feature of the drone system is the positioning of the IMP sampler high enough above the propellers, which ensures that the sampled air remains free from propeller-induced turbulence, commonly known as downwash. The drone was flown at a fixed altitude of 10

meters during each sampling interval, a height selected to prevent propeller downwash from disturbing the switchgrass canopy during stable hovering.

Due to drone battery limitations and the need for safe flight and landing operations, each sampling interval was restricted to 10 minutes. The IMP unit on the drone operated at the same volumetric sampling rate as its ground-based counterpart (0.6 L/min). However, because the drone sampler was only flown for 10 minutes per flight, its total sampling capacity was significantly lower than the ground-based IMP units, which collected for 30 to 90 minutes during the field campaign. Hereafter, the IMP-equipped drone system is referred to as "DRN" throughout the manuscript.

3.2.3 Meteorological Data

A weather station was installed near the field to collect meteorological data throughout each sampling day. The station consisted of a Campbell Scientific CSAT3 three-dimensional sonic anemometer, mounted at a height of 1.5 m above ground level, which measured high-resolution wind velocity in three dimensions and sonic temperature at a frequency of 10 Hz. In addition, a Campbell Scientific HMP45C probe recorded temperature and relative humidity every 30 seconds. Meteorological data were recorded with the Campbell Scientific CR3000 datalogger. To minimize interference from the tripod pole, the sonic anemometer arm was positioned perpendicular to the anticipated dominant wind direction before each collection day. The wind velocities in the u and v directions (relative to the sonic anemometer arm) were then rotated into the cardinal coordinate system for analysis.

3.2.4 Processing of pollen samples

Sample preparation

Filters from the ED samplers were processed as shown in Figure 3.4A. Briefly, the 125mm "ED" collection filters were removed with forceps and transferred to separate 150 mm petri dishes (Fisher #FB0875714) in the field immediately following each sampling period. For each filter, 25 mL of 25% EtOH was added in the petri dish, the filter was gently agitated with a sterile cell spreader, and then rinsed a total of 8 times. Each rinsate was transferred by a pipettor to a vacuum filtration unit (Thomas Scientific #300-4100) containing a 47 mm Isopore polycarbonate 10μm filter (Millipore Sigma #TCTP04700). The sample was cleared through the filter using the vacuum from a hand pipetting bulb. Using forceps, the Isopore filter was then moved to a 60mm petri dish (Genesee 32-105) and rinsed 6 times with 2 mL 25% EtOH. The resulting rinsate was transferred to an Ultrafree 5 μm PVDF centrifugal filtration tube (Millipore Sigma UFC40SV25) and centrifuged for 2 minutes at 2,500 rpm in a swinging bucket centrifuge (IEC Clinical Centrifuge). The concentrated sample was then resuspended from the 5μm filter surface with 200μL 25% EtOH and moved to a 1.5 mL Eppendorf tube and stored at 4 °C for further analysis

Filters from the FRM sampler were processed as shown in Figure 3.4B. Briefly, the Isopore filter was removed from the FRM unit sampling cartridge using forceps and transferred to a 60 mm petri dish (Genesee 32-105) in the field immediately following each sampling period. For processing the sample, the Isopore filter was then rinsed 6 times with 2mL 25% EtOH, and the resulting rinsate was moved to an Ultrafree 5m PVDF centrifugal filtration tube (and centrifuged for 2 minutes at 2,500 rpm in a swinging bucket centrifuge (IEC Clinical Centrifuge). The concentrated sample was then resuspended from the 5m filter surface with 200 L of 25% EtOH and moved to a 1.5 mL Eppendorf tube and stored at 4 °C for further

analysis.

The fluid from the IMP and DRN samplers was processed as shown in Figure 3.4C. Samples from the IMPs and DRN were transferred by pipette to an Ultrafree 5μm PVDF centrifugal filtration tube and centrifuged for 2 minutes at 2,500 rpm in a swinging bucket centrifuge (IEC Clinical Centrifuge). The concentrated sample was then resuspended from the 5 μm filter surface with 200 μl 25 EtOH and transferred to a 1.5 mL Eppendorf tube and stored at 4 °C for further analysis.

Pollen counting

Switchgrass pollen was counted in each concentrated sample by pipetting the samples into individual wells of a 96-well plate (Grenier Bio One 7000124). Samples were allowed to sediment for 15 minutes and then were observed using an Olympus CKX53 inverted microscope equipped with the Olympus EP50 digital camera and associated software. Following the quantification of the switchgrass pollen in each of the tubes, the samples were transferred back into their respective 1.5 mL Eppendorf storage tubes and held at 4 °C for transport and further analysis.

3.2.5 Atmospheric dispersal modeling

Meteorological Inputs

Atmospheric dispersal simulations for each sampling interval are driven using time-averaged wind statistics collected during that interval. Most sampling intervals occurred under low-wind conditions ($<$ 2 m/s), characterized by meandering winds with frequent directional shifts and intermittent lulls in wind speed. To better capture dispersal dynamics under these

conditions, the wind data was processed using different averaging window sizes. Specifically, for the 45-minute sampling intervals on August 2-3, 2022, the following averaging windows were used: 45 one-minute averaging windows, 9 five-minute averaging windows, and 1 full 45-minute averaging window. This approach allowed for assessing how different temporal resolutions of wind averaging influenced the accuracy of dispersal simulations.

To compute turbulence statistics for each averaging window, the average downwind direction was determined and the wind velocity data was rotated into the downwind (u)and crosswind (v) coordinate system. For each sampling interval, the means, covariances, and variances were computed for these wind velocity components, as well as temperature. Mean temperature was computed from sonic temperature using the method described in [143] and the relative humidity values. Heat flux was estimated from sonic temperature and relative humidity using the Bowen ratio method from [143]. The Bowen ratio was determined using the simplified method of [144], which requires only mean temperature and relative humidity. These turbulence statistics provided the necessary meteorological inputs for the dispersal simulation in each interval, specifically friction velocity (u_*) and the Monin-Obukhov length (L).

Pollen dispersal simulations

To simulate switchgrass pollen dispersal, we use the surface-layer Lagrangian Stochastic (LS) model described in [31] and expanded to three-dimensional transport in [75]. The LS model framework is based on Brownian motion theory, modeling turbulent diffusion by simulating the trajectories of thousands of particles through the air as random walks through the atmosphere. The movement of each particle is governed by turbulent wind statistics, and the ensemble average of these trajectories provides estimates of pollen concentration at any given location within the simulation domain.

LS models require turbulence wind statistics to be specified at every point in the simulation domain, including mean velocities, variances, and covariances of the wind components. Under the assumptions of stationarity and horizontal homogeneity, these wind field statistics remain constant over time within each averaging window but vary with height. To account for this height dependence, boundary layer scaling techniques are applied to generate vertical wind profiles based on measurable surface-level parameters, specifically the friction velocity (u_*) and the Monin-Obukhov length (L). The full model formulation and wind statistics used in the simulations are included in the code provided online.

These two parameters (u_* and L) were computed from the time-averaged meteorological measurements for each 45-minute sampling interval, using three different averaging approaches: one 45-minute averaging window; nine 5-minute averaging windows; and forty-five 1-minute averaging windows. A separate simulation was conducted for each averaging window in a 45-minute sampling interval, using the computed average wind direction, friction velocity (u_*), and Monin-Obukhov length (L). The resulting concentration fields from these simulations were then averaged to generate a single mean concentration field for each 45-minute sampling interval.

In each simulation, 100,000 particles (representing switchgrass pollen) are released from a point source at the center of the field at a height of 2 m, which approximates the height of most of the switchgrass panicles in the field experiment. Particles were removed from the simulation domain when they: traveled more than 50 m laterally from the source; rose above 100 m above ground level; or fell below 0.1 m above ground level. To simplify the simulation, pollen dispersion was modeled as if it occurred over a flat, rough surface, with estimated surface roughness of 0.01 m, consistent with values reported for level grassy plains and prairie in [145]. The settling velocity was estimated as 0.0371 m/s, based on an observed switchgrass pollen diameter of $\sim$35 μm, using Stokes law. Since switchgrass pollen was observed to be

nearly spherical, Stokes law provides a reliable approximation of its settling velocity.

Concentration estimation and source emission rate calculation

The pollen concentration estimation procedure in this study follows the approach described by [146] for a stationary LS model with a constant source. Pollen concentration is estimated by tracking the amount of time particles spend in each grid box, normalized by the total number of particles released (N_{p}) and the volume of the grid box (V_{box} = 1 m × 1 m × 1 m), and then multiplied by the modeled emission rate (Q). Specifically, the concentration at a given grid point (i, j, k) is calculated as,

$$C(i,j,k) = Q\frac{1}{V_{\mathrm{box}}}\frac{1}{N_{\mathrm{p}}}\sum_{n=1}^{N_{\mathrm{p}}} T_n(i,j,k), \tag{3.1}$$

where $T_n(i, j, k)$ represents the time particle n spends in the given grid box.

We employed a model-measurement fusion approach described in [31] to estimate pollen emission rate and concentration. For each sampling interval, the modeled pollen concentration at each grid point in the simulation domain is first computed under the arbitrary assumption of a constant release rate at the center of the field of $Q_{\mathrm{model}} = 1$ particle per second. This yields a modeled relative concentration, which is proportional to the actual concentration at every point in the domain. To estimate the actual emission rate (pollen flux from the field), the ratio of the measured pollen concentration at each of the six ED samplers to the modeled relative concentration at the corresponding locations in the simulation domain was computed and used to update the value of Q. To obtain a single emission rate estimate for each sampling interval, the computed emission rates corresponding to calculations based on each of the six ED samplers were averaged. This estimated true emission rate was then used to update the modeled relative concentration to predict the actual concentration at every

point within the simulation domain.

To investigate dispersal at greater distances—up to 1 km from the source—the same modeling procedure is conducted but with a coarser grid resolution of $V_{\text{box}} = 3\text{ m} \times 3\text{ m} \times 1\text{ m}$ and an extended simulation domain covering 1000 m × 1000 m × 100 m. This coarser grid was selected to balance computational efficiency while maintaining consistency with the finer-resolution near-source grid.

To generate 2D dispersal kernels, which represent pollen concentration as a function of distance, concentrations at equal radial distances from the source are averaged, yielding average concentration as a function of radial distance from the source.

3.3 Results

Three field campaigns were conducted over the course of two calendar years (2021 and 2022) to sample airborne pollen around two blocks of transgenic switchgrass.

3.3.1 Field Experiments

First campaign (October 7-8, 2021)

The first field campaign took place on October 7-8, 2021. At this time, the pANIC10A and PSYBIN1a plants (Fig. 3.1) were only 11 and 6 weeks old, respectively, and the youngest 200 pANIC10A plants had not yet been planted. The following ground based samplers were placed radially around the field at distances of 0, 5, 15, and 25 meters, as shown in figure 3.2: three ED samplers, three IMPs, and one FRM sampler. These ground-based samplers operated in 30-minute intervals, while the drone-based sampler (DRN) was flown

at a 10-meter height for a duration of 10 minutes per sampling interval. However, due to technical difficulties, the drone sampler was only deployed on October 7th, and was not used on October 8th.

IMP A was placed at the center of the field (the midpoint of field 1 and 2) to estimate pollen emission rate, but an insufficient amount of pollen was collected—at most two pollen grains in each sampling interval, and often zero—which was not statistically significant to estimate the pollen emission rate. IMP B and ED C were co-located to verify alignment of their concentration estimates. However, due to their vastly different sampling rates and the low pollen numbers collected, direct comparisons were not feasible. On October 8th, slightly more pollen was captured, particularly by ED C, but overall pollen collection remained low. Impingers showed an increase in measured concentrations later in the afternoon, though data remained sparse. The IMP, FRM, and DRN samplers captured negligible pollen amounts.

Pollen concentrations for each sampling interval and sampler are presented in Table 3.1. On October 7th, too few pollen grains were collected for meaningful analysis. Concentrations marked with a dagger†superscript denote cases where only 1-2 pollen grains were sampled. On October 7th, all samplers captured 2 or fewer pollen grains. On October 8th, 1-5 grains were collected per sampler. IMP B did not capture any pollen, despite being placed alongside ED C, likely due to the overall low pollen counts.

Second campaign (October 20-21, 2021)

During the second field campaign on October 20-21, 2021, we increased ground-based sampling intervals to 90 minutes to compensate for the low pollen counts observed during the previous campaign's 30-minute sampling intervals. To improve pollen capture, all samplers were moved within 15 meters of the field. The drone was still flown for 10 minutes per

sampling interval. On October 21st, heavy rainfall forced us to shorten the final sampling interval of the day. Despite the longer sampling durations and closer sampler placement, pollen counts remained negligible (at most two pollen grains captured per interval) throughout this campaign. See Table 3.1 for details.

Third campaign (August 2-3, 2022)

The third and final campaign took place on August 2-3, 2022. Field and sampler placements during this campaign are shown in Figure 3.1A and B, respectively. The oldest pANIC10A plants were at peak pollen production during this campaign, but the PSYBIN1a plants had not yet begun releasing pollen. The youngest pANIC10A field, planted only six weeks prior, was too immature to release pollen. Given that ED samplers were the most effective in previous campaigns, their number was increased from three to six. In anticipation of prevailing winds directed toward north-northeast, samplers were primarily aligned along the x and y axes in Figures 3.1 and 3.2. The drone sampler was again flown for 10 minutes at 10 meters during all sampling intervals. ED D and IMP A were co-located to compare concentration measurements. IMP B was placed at the field center to estimate the pollen emission rate, if sufficient pollen was collected. To address diurnal trends, sampling intervals were kept consistent across both collection days. All ground-based samplers were operated for 45-minute sampling intervals. ED samplers collected significantly more pollen than in previous campaigns. A clear diurnal pattern emerged, with concentrations peaking between 2:00-2:45 PM time on both days. Concentrations began increasing around 1:00 PM, peaked at 2:00 PM, then declined after 3:00 PM. See Table 3.1 for details.

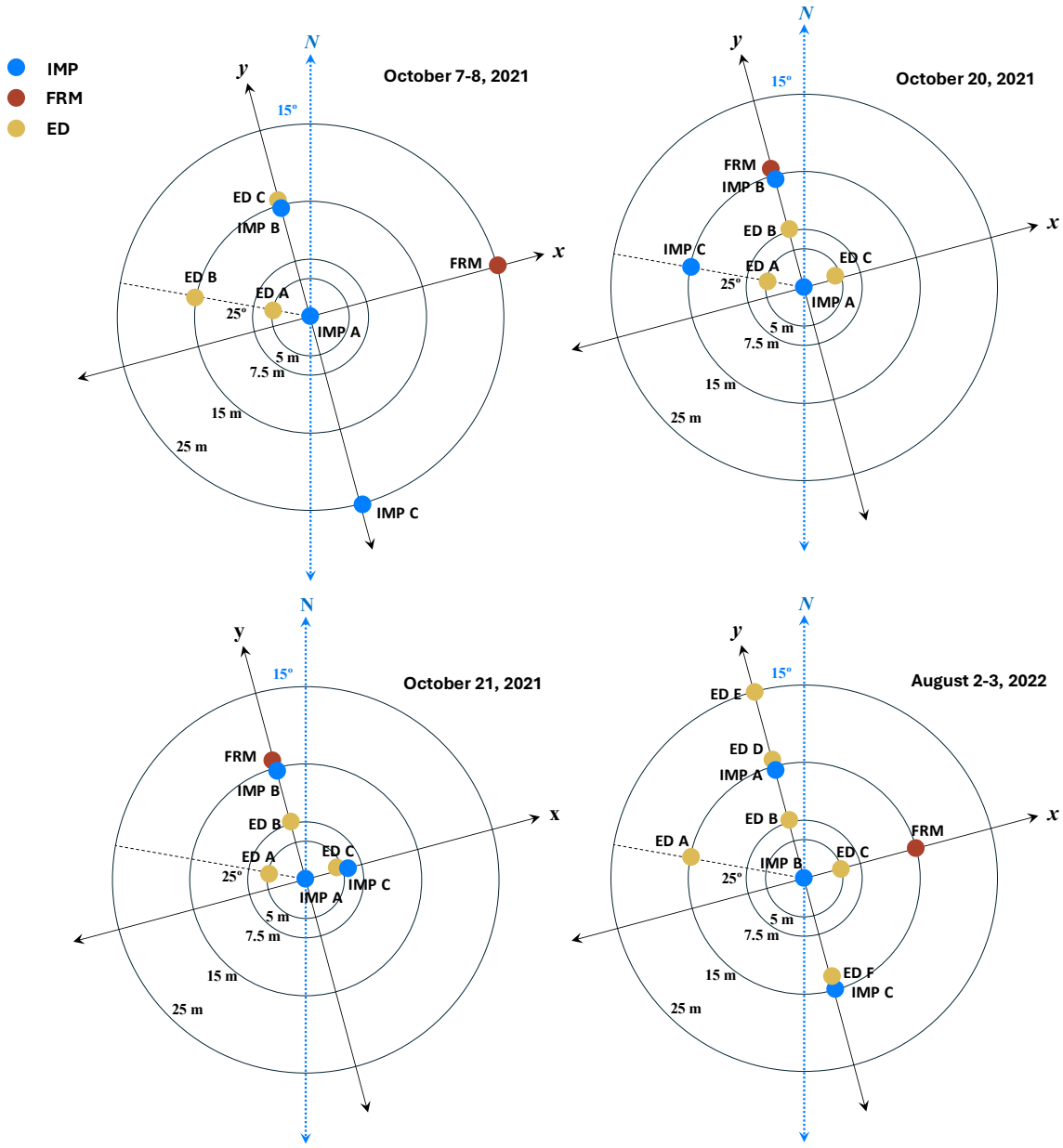


Figure 3.2: Ground-based sampler locations for each collection day. Yellow, blue, and red denotes placement of ED, IMP, and FRM samplers, respectively, positioned at radii of 5, 7.5, 15, and 25 meters from the center of field.

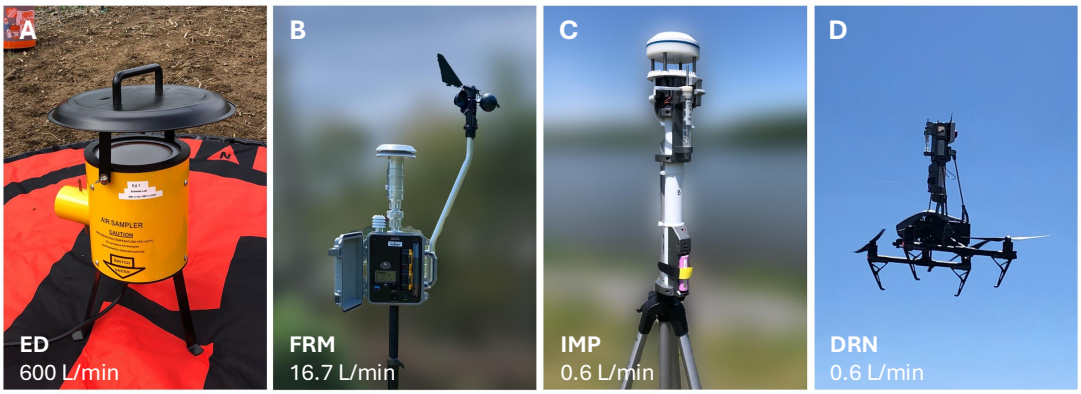


Figure 3.3: All sampler units used during the field campaigns. (A) The ED[1], a ground-based high-volume sampler (600 L/min). (B) The FRM, a ground-based medium-volume volumetric sampler (16.7 L/min). (C) The IMP, an impinger-based ground-based low-volume sampler (0.6 L/min). (D) The DRN, a drone-based sampler flown at a height of 10 meters above ground-level (0.6 L/min).

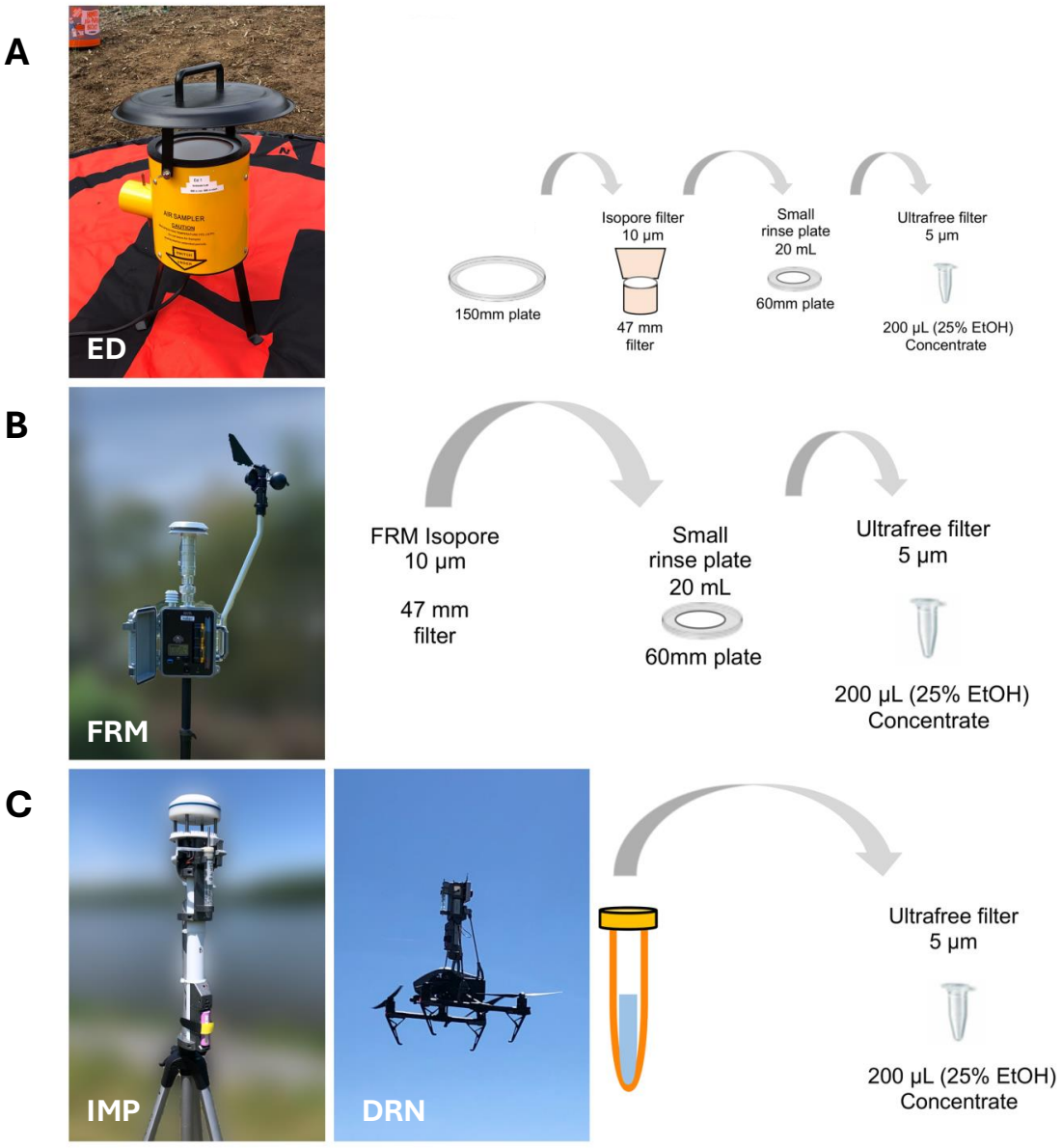


Figure 3.4: Flow charts showing the processing of the filters from (A) the ED samplers, (B) the FRM sampler, and (C) the IMP and DRN samplers.

		ED						FRM	IMP			DRN**
Date	Time	A	B	C	D	E	F		A	B	C	
10/7/21	11:30-12:00	0	0.17	0	-	-	-	0	0	-	0	0
	12:35-13:05	0	0	0.06†	-	-	-	0	0	0	0	0
	13:45-14:15	0.06†	0	0	-	-	-	0	0	0	0	0
	14:45-15:15	0.11†	0	0.11†	-	-	-	2.00†	0	0	0	0
	15:50-16:20	0.17	0.06†	0.11†	-	-	-	0	0	0	0	0
10/8/21	12:00-12:30	0	0	0.33	-	-	-	2.00†	0	0	0	-
	13:00-13:30	0.17	0.17	0.06†	-	-	-	0	0	0	0	-
	14:00-14:30	0	0	0.22	-	-	-	0	111.11†	0	0	-
	15:00-15:30	0.11†	0.17	0.33	-	-	-	0	0	0	0	-
10/20/21	14:00-16:00*	0.02†	0.02†	0.02†	-	-	-	0.67†	0	0	0	0
	16:30-18:15*	0.02†	0.02†	0.02†	-	-	-	0	0	0	0	0
10/21/21	12:00-13:45*	0	0	0.02†	-	-	-	0	0	0	0	0
	14:00-14:50	0	0	0.03†	-	-	-	0	0	0	0	0
8/2/22	10:00-10:45	0.37	0.48	2.48	0.15	0.04†	0.04†	0	0	0	0	0
	11:00-11:45	0.11	0.89	0.59	0.04†	0.04†	0.04†	0	0	0	0	0
	12:00-12:45	0.07†	3.19	1.04	0.37	0.22	0.11	0	0	0	0	0
	13:00-13:45	0.22	4.30	31.78	0.33	0.26	1.89	0	0	0	0	-
	14:00-14:45	0.37	41.44	19.74	4.67	1.00	1.59	15.97	74.07†	37.04†	0	166.67†
	15:00-15:45	0.26	5.41	5.85	0.15	0.11	0.22	3.99	0	0	37.04†	0
8/3/22	10:00-10:45	0.04†	0.15	0.33	0	-	0.07†	0	0	37.04†	0	333.33†
	11:00-11:45	0	1.74	0.30	0.07†	0	0.15	0	0	0	0	0
	12:00-12:45	0	0.48	0.85	0.04†	0.15	0.07†	0	0	0	0	0
	13:00-13:45	0.15	6.22	17.89	0.74	0.04†	0.11	1.33†	0	37.04†	37.04†	0
	14:00-14:45	15.11	23.44	29.26	9.67	1.70	3.19	10.65	0	444.44	74.07†	500.00
	15:00-15:45	0.78	5.19	9.07	1.44	-	0.48	13.31	111.11	185.19	37.04†	0

Table 3.1: Pollen concentrations (pollen/m^3) sampled during specific time periods, organized by date and sampler type.
†Only 1-2 pollen grains were sampled.
*sampling times for ED and FRM were limited to two 45 minute intervals with a break in between to prevent overheating, resulting in 90 minutes of total sampling time. Impingers sampled for the entire time.
**drone was sampling for 10 minutes in every interval due to battery limitations

3.3.2 Orange-Fluorescent Protein Expression

The primary source of pollen in the field experiments came from the first planted block of GE switchgrass, which contained 100 plants expressing OFP under the control of a maize ubiquitin promoter (PvUBI1). As shown in Figure 3.5, the OFP signal in pollen grains from these transgenic plants was difficult to distinguish from wildtype (WT) pollen exposed to the same OFP-inducing wavelength of light. In contrast, the OFP signal in pollen from the later planting of GE switchgrass (expressing OFP under a maize pollen-specific promoter, Zm13) was much stronger and easily distinguishable from WT pollen. However, these Zm13-expressing plants were smaller and did not produce sufficient mature panicles in time for the field experiments, limiting their contribution to the study.

3.3.3 Modeling Results

For the dispersal modeling, we focused on the third field campaign (August 2-3, 2022), as sufficient pollen was captured on both days to allow for concentration measurements from the ED samplers. All sampling intervals during these days were 45 minutes long. Sampling occurred at consistent times across both days, facilitating comparisons and enabling the identification of diurnal trends.

Near-source concentration and emission rate estimation

Dispersal simulations more accurately capture changing wind directions and pollen concentrations when shorter averaging periods are used for each sampling interval. Figure 3.6 presents ground-level relative concentration contours and downwind wind roses for the August 2nd, 2:00-2:45 PM sampling interval, simulated using 1-minute, 5-minute, and 45-minute

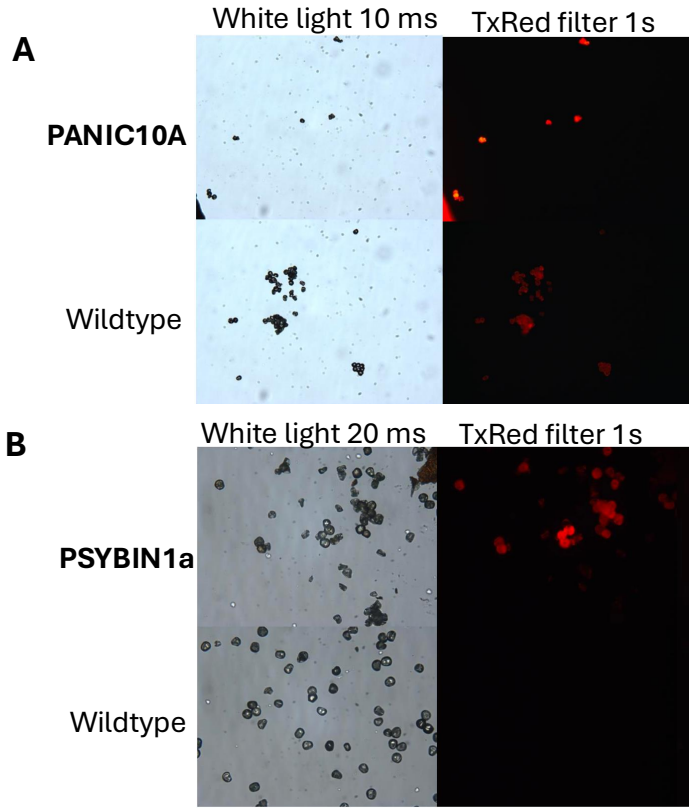


Figure 3.5: Comparison of OFP signal between GE switchgrass pollen and wildtype. (A) Pollen from the pANIC10A strain (OFP expression throughout the entire plant) is not easily distinguishable from wildtype pollen under OFP-inducing light. (B) Pollen from the PSYBIN1a strain (OFP expression restricted to pollen) exhibits a strong, highly distinguishable OFP signal compared to wildtype pollen.

averaging windows. This interval corresponded to the highest measured pollen concentrations by the ED samplers. The wind rose in Figure 3.6A was generated using wind data collected at a sampling frequency of 10 Hz. The wind roses in Figure 3.6B, C, and D were created using 1-minute, 5-minute, and 45-minute time-averaged wind data, respectively. Wind roses display the downwind direction. Yellow circles indicate the ED sampler locations, with their size proportional to pollen counts at each site.

Comparison of averaging windows. The 45-minute plot (Figure 3.6D), based on a single LS simulation in the average downwind direction, fails to capture wind variability and lateral

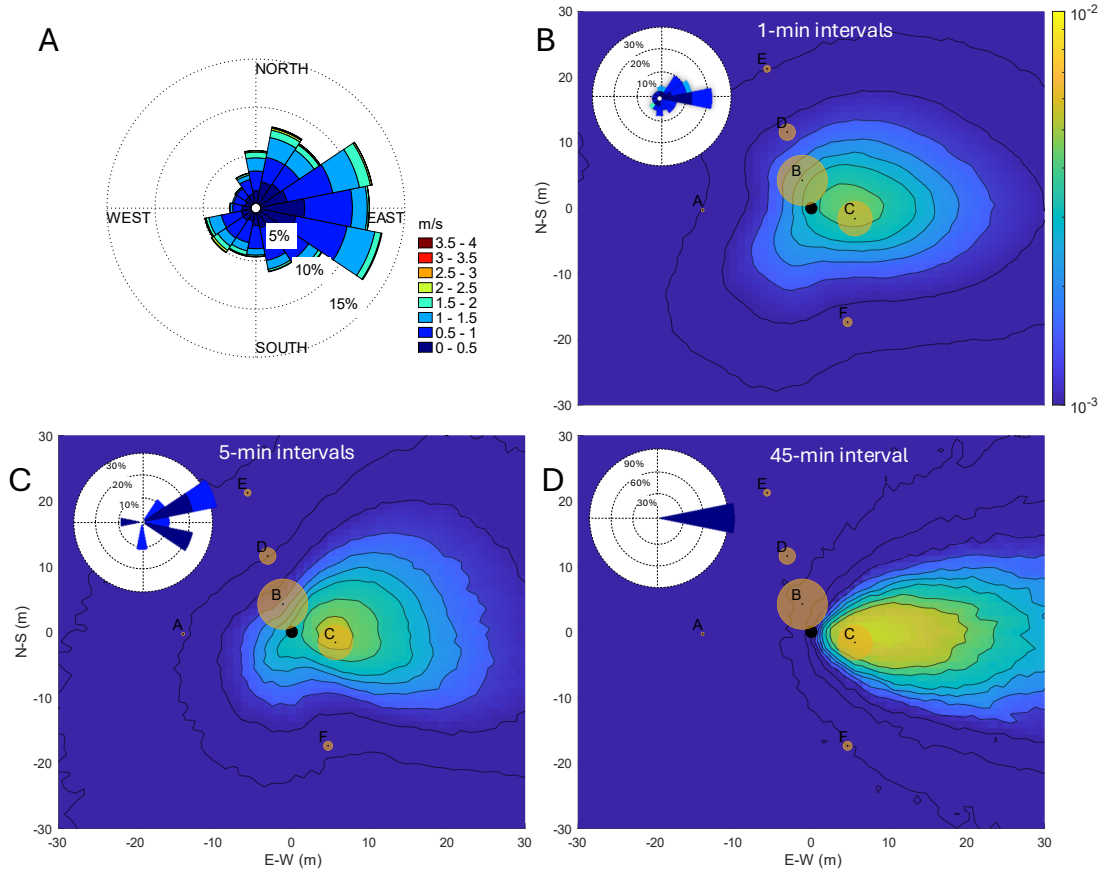


Figure 3.6: Relative ground-level concentration contours for the August 2nd, 2:00-2:45 PM sampling interval. (A) Wind rose for this sampling interval, indicating the downwind direction. (B) Contour plot averaging forty-five 1-minute simulations. (C) Contour plot averaging nine 5-minute simulations. (D) Contour plot generated using a single 45-minute simulation. Yellow circles indicate the ED sampler locations, with their size proportional to pollen counts at each site.

pollen spread, missing high pollen counts at ED sampler B due to a single eastward-directed plume. The 5-minute plot (Figure 3.6C), which averages nine LS simulations, shows some directional variation but lacks the detail seen in the 1-minute plot. The 1-minute plot (Figure 3.6B) provides the most accurate representation of dispersal dynamics. However, all three simulations share a common discrepancy: peak concentrations appear a few meters from the point source, indicating lateral transport before deposition. A field source, rather than a point source, may better address these inconsistencies, at the cost of further model

complexity.

Emission rate and diurnal pattern. The computed pollen emission rate from the field exhibits a clear diurnal trend. Figure 3.7 presents the mean, minimum, and maximum non-zero computed emission rates for each sampling interval on August 2-3, 2022 during the third field campaign. Emission rates were computed from Equation (4.2), based on the ratio of measured to modeled concentrations. Samplers with zero measured or modeled concentrations were excluded to prevent infinite or zero emission rate estimates. Emission rate calculations were performed for 1-minute, 5-minute, and 45-minute averaging windows. The range of estimated emission rates decreases with smaller averaging windows, indicating greater precision with shorter averaging periods. The pollen emission rate increased by approximately three orders of magnitude between 10 AM and 2 PM. Note that the emission rate is shown on a log scale in Figure 3.7. The log-transformed emission rate estimate is positively correlated with the horizontal velocity magnitude (Pearsons R = 0.73, P = 0.01), temperature (R = 0.71, P = 0.01), and vapor pressure deficit (R = 0.69, P = 0.01), while negatively correlated with relative humidity (R = -0.63, P = 0.02). These results indicate that higher pollen emissions occur under conditions of higher wind speed, temperature, and vapor pressure deficit, while increased relative humidity reduces pollen release.

Modeled concentration predictions improve dramatically when changing wind conditions are incorporated into the simulations. Figure 3.8 compares modeled concentrations, computed by multiplying the estimated emission rate by the relative concentration, with measured concentrations derived from ED sampler pollen counts. Although the figures directly compare modeled and measured concentrations, they are not intended as a formal model validation, as the measured concentrations were directly used to compute emission rate and modeled concentration (see Section 2.5.3). Instead, they highlight the substantial improvement in model performance as the averaging window is reduced. The Pearson's R-value increases

from 0.19 for 45-minute windows to 0.71 for 5-minute windows to 0.84 for 1-minute windows. While reducing the averaging window from 45 minutes to 5 minutes requires nine times the computational power, it yields a 270% improvement in model performance (as measured by the R-value). In contrast, refining the resolution further from 5-minute to 1-minute windows results in only a 20% increase, suggesting that shorter windows may not always be computationally worthwhile beyond a certain threshold.

Estimating sensor capabilities in the far field

The intake rates of the samplers used in this study differ by multiple orders of magnitude, with 0.6 L/min for the low-volume IMP and DRN, 16.7 L/min for the medium-volume FRM, and 600 L/min for the high-volume ED. This disparity highlights the coarseness of concentration measurements for the low-volume IMP and DRN samplers, particularly when considering error estimates based on the Poisson distribution. The Poisson distribution models the error of discrete event counts, such as the number of pollen grains collected, as $\sqrt{N}$, where N is the number of pollen grains captured by a sampler [147, 148]. Lower intake rates necessitate proportionally higher aerial concentrations, and consequently, greater pollen emission rates for detection, leading to significantly higher measurement error compared to high-volume samplers.

To determine optimal sampler placement based on varying field emission rates, we combined sampler intake rates with long-distance dispersal simulations. Figure 3.9 plots the maximum distances at which each sampler could be placed to collect at least 100 pollen grains ($\pm$10) within a 45-minute sampling interval, as a function of the pollen emission rate. These estimates were generated using the long-distance 2D relative concentration dispersal kernels, computed separately for ground-based samplers (IMP, ED, FRM) at ground level and the drone-based sampler (DRN) at 10 m above ground level. Although the drone was only flown

for 10 minutes per sampling interval in our field campaigns, we evaluated its performance over 45-minute intervals to account for potential future battery-life improvements or wire-tether power modifications. The threshold of 100 pollen grains corresponds to different effective concentrations across sampler types, with 3.7×10^3 pollen/m^3 for IMP and DRN, 133 pollen/m^3 for FRM, and 3.7 pollen/m^3 for ED. This threshold was selected as it ensures a $\pm10\%$ error in concentration estimates when modeled with the Poisson distribution.

Figure 3.9 shows that, to collect at least 100 pollen grains, the ED samplers require 1.5 orders of magnitude less emission rate than the FRM and 3 orders of magnitude less than the IMP, across all distances from the field. This advantage allows ED samplers to be placed significantly farther from the field compared to IMP and FRM samplers. The DRN requires an even greater emission rate, due to lower concentrations at 10 meters altitude. However, at approximately 200 meters downwind, the DRN and IMP detection capabilities converge, as vertical concentration gradients become less pronounced further from the source.

The gray vertical lines in Figure 3.9 represent computed emission rate values at each ED sampler during our field campaign, which are also shown in Figure 3.7. These results indicate that only the ED samplers had a reasonable chance of collecting at least 100 particles during some sampling intervals. The remaining samplers were largely ineffective at detecting airborne pollen given the small emission rates observed in this study, even at close proximity to the source field. For a larger source, with an emission rate of 10^6 pollen/s (approximately 100 times larger than ours), the feasible placement of samplers would improve substantially. Under such conditions, impingers could be placed up to 10 meters from the field, the FRM samplers could be placed up to 100 meters from the field, and ED samplers could be placed up to 200 meters away. These adjustments would allow each sampler type to collect at least 100 pollen grains within a 45-minute sampling interval, enhancing measurement reliability and reducing uncertainty.

3.4 Discussion

In these field experiments, airborne pollen from two different strains of OFP-expressing switchgrass plants were captured and analyzed using the method described in [140]. The PSYBIN1a strain exhibited strong fluorescence compared to wild-type pollen, whereas the pANIC10A strain was difficult to distinguish by fluorescence alone. However, since the PSYBIN1a plants did not release sufficient pollen during any of the field experiments, OFP expression was not used for pollen identification in this study. If we had successfully captured PSYBIN1a pollen from the younger field, it would have expedited the counting and sampling process. Despite this limitation, the study serves as a proof-of-concept that fluorescence tagging can be a valuable tool for pollen tracking. Though prior studies have tracked the movement of GE pollen in the atmosphere [11], to our knowledge, this is the first detailed study to incorporate aerial and ground-based volumetric samplers to track the movement of GE pollen from a known source. Fluorescent tagging presents a unique opportunity to trace pollen dispersal and track its movement over long distances. Understanding switchgrass gene flow is particularly relevant as biofuel production increases, helping to mitigate ecological risks posed by invasive strains and unintended cross-pollination between transgenic varieties [19]. Additionally, fluorescence-tagged pollen could facilitate rapid and automated counting using instruments such as the Helmut Hund BAA500, Plair Rapid-E, or WIBS-4 [133, 149], which is of particular interest for allergen monitoring.

During the final campaign on August 2-3, 2022, a distinct diurnal pattern emerged in both measured pollen concentrations and computed pollen emission rate. Pollen emission rate increased steadily after the first sampling interval at 10 AM, peaked at 2 PM on both days, and declined during the final sampling interval at 3 PM. This diurnal pattern was consistent across both days and correlated with increasing wind velocity and temperature,

as well as decreasing relative humidity. Such diurnal pollen release patterns are common in wind-pollinated species, where anther dehiscence is driven by drying conditions such as low humidity and rising temperatures [23]. Similar trends have been observed in previous switchgrass field studies, where peak pollen concentrations occurred in the late morning and early afternoon, followed by a decline around 3 PM [126]. Comparable findings in maize have linked pollen release patterns to increasing vapor pressure deficit [51] and decreasing relative humidity combined with rising wind velocity [55]. This information can be used to better predict peak allergen concentrations and improve accuracy of large-scale air pollution models.

Throughout all sampling intervals and field campaigns, we observed very low wind velocities (< 2 m/s) and frequent shifts in wind direction. Under these meandering wind conditions, particle dispersal is primarily controlled by wind direction shifts rather than turbulence [150, 151]. Standard dispersal models that assume a dominant downwind direction fail to account for this effect, often producing overly narrow plumes that underestimate lateral spread. This limitation is particularly characteristic of LS models [51, 152] and Gaussian plume models [153] which require a single downwind direction. Even more advanced modeling approaches, such as Large Eddy Simulations, do not fully incorporate changing wind directions [154]. To address this, we reduced the averaging window for wind statistics from 45 minutes to 5 and 1 minute, then ran dispersal models for each of these smaller intervals and combined the resulting plumes. This approach dramatically improved the fit between modeled and measured concentrations, enhancing emission rate estimates. Similar techniques have been applied in Gaussian plume modeling with 2-minute intervals, yielding significantly better agreement with measured data [153]. Anfossi et al. (1990) [152] also emphasized the importance of using short averaging windows for dispersal modeling, recommending intervals of only a few minutes. A maize dispersal study similarly attributed discrepancies between

measured and modeled concentrations to wind direction variability and assumptions of a dominant wind direction [51]. Future large-scale pollen forecasting and bio-confinement strategies should consider meandering wind conditions, which are not currently accounted for in large-scale models [154].

The highest concentration measurements in this study came from the high-volume ED samplers. The pollen source size—100 plants releasing pollen—was exceedingly small compared to previous dispersal experiments in switchgrass, which involved 3,200 plants [126], as well as studies in maize [51, 53, 55]. High-volume ED samplers performed best under these small-source conditions, capturing spatial variations in concentration and diurnal patterns. To the best of our knowledge, this is the first pollen-trapping field study to utilize these samplers. Their volumetric flow rate of 600 L/min is 60 times greater than that of the traditional Hirst-type samplers, which operate at 10 L/min and are comparable to the FRM sampler used in this study. Due to the small pollen source, the FRM sampler did not produce usable data. As the ED samplers were originally designed for educational purposes, they are inexpensive and lack pre-programming and other advanced features found in commercial volumetric samplers. However, their simplicity and affordability make them easily deployable, and they have strong potential for measuring concentrations from small sources and capturing high-resolution pollen dispersal patterns even in small fields.

A novel impinger-type particle sampler (IMP and DRN) was used in this study to collect pollen, marking the first application of this integrated system for pollen tracking. While previously employed for airborne microbial sampling[24, 25], this study extends its use to pollen dispersal. The IMP and DRN samplers successfully collected pollen in the field, demonstrating their feasibility for tracking pollen movement. However, due to the small source size, limited pollen production, and the relatively low sampling rate of 0.6 L/min, the collected pollen quantities were insufficient for reliable concentration estimates. The dif-

fering flow rates between the ED and IMP samplers further complicate direct comparisons, as impinger samplers inherently capture fewer particles at high concentrations due to their small intake volumes. The IMP system would likely perform more effectively when sampling from much larger sources, at least 100 times the size of the field used in our study (Figure 3.9). Similarly, the drone-mounted sampler, operating at 10 m AGL, would require a significantly larger pollen source for effective deployment at further distances and altitudes. Nevertheless, the drone platform remains a valuable tool for aerobiological research, offering future opportunities for prolonged and spatially resolved sampling, particularly when paired with higher-volume sampling technologies, including those incorporating filter-based collection systems. Moreover, impinger samplers, which preserve particles in liquid, could prove especially useful for future viability and molecular studies.

A primary constraint in regulated transgenic pollen dispersal studies is the feasible scale of the pollen source. From an agricultural perspective, a 100-plant source is small relative to other agricultural dispersal studies; however, in the context of permitted flowering GE pollen dispersal, it is substantial because it requires specialized propagation, isolated siting, and a stringent regulatory permitting. Our experimental design therefore prioritized a small, but well-contained OFP-tagged source, and we structured the sampling campaign to extract the most robust insights that this rare setup could support. The value of the present study is therefore not that it reproduces a large agricultural pollen release, but that it establishes what is measurable and how to model it under realistic constraints that are intrinsic to transgenic field experimentation. While future studies would benefit greatly from a larger pollen source, the small source and low-wind conditions allowed us to stress test our sampling and modeling methodology.

With unlimited resources, the natural next step would be a scaled up version of this same experiment focused on validation: a large, well-established source containing only PSYBIN1a

to maximize the OFP signal, multi-year sampling on the same dates and times to quantify repeatability of diurnal emission patterns, and colocated Hirst-type gold-standard samplers deployed alongside the novel samplers to provide independent concentration estimates for model validation and rigorous sampler inter-comparison. A substantially larger source would allow meaningful comparisons between low- and high-volume samplers, with one set used to estimate the particle release rate and another for validating modeled concentrations. It would also enable more effective use of impinger-type samplers (IMP and DRN), which could preserve pollen for downstream viability studies, although isolated siting requirements for transgenic work would continue to pose limits for long-distance tracking.

Focusing solely on PSYBIN1a switchgrass, with its stronger OFP fluorescence in pollen, could further enhance tracking accuracy via automatic fluorescence-based quantification. As shown in Figure 3.9, these methods could enable pollen detection up to 1 km or even tens of kilometers from the source. This fluorescence tagging technique could also be transferable to other crops of interest. For instance, hemp is known to produce copious amounts of pollen capable of long-distance dispersal, and its monitoring is increasingly relevant [123]. If transgenic hemp lines become available, similar fluorescence-based tracking methods could be applied to study its pollen movement and gene flow in detail.

3.5 Conclusion

Three field campaigns were conducted to measure pollen concentrations around a small field of genetically modified switchgrass, utilizing both traditional and novel sampler types. The experiments also included a drone-mounted sampler, demonstrating the feasibility of airborne pollen sampling at 10 meters above the field as a proof-of-concept. Despite the exceedingly small source size, the high-volume ED samplers successfully collected sufficient

pollen to analyze spatial variations in concentration and identify diurnal release patterns. This study evaluated the effectiveness of different sampler types for pollen collection under varying conditions. Among the three field campaigns, only the final campaign on August 2-3, 2022, produced sufficient concentration data for detailed analysis and modeling. During this campaign, a clear diurnal pattern was observed in the pollen concentration and, consequently, in the calculated emission rate. Persistent low-wind meandering conditions were recorded throughout the campaign, and reducing the averaging window for simulations significantly improved pollen emission rate estimations by better incorporating shifting wind directions. This study highlights the potential for drone-based pollen sampling and fluorescence-based GMO pollen tracking. The findings provide insight into the effectiveness of different sensor types with respect to source strength and sampling distance, advancing the understanding of pollen dispersal dynamics and measurement techniques. These results have important implications for allergen monitoring, cross-pollination risk assessment, and broader bioaerosol surveillance strategies.

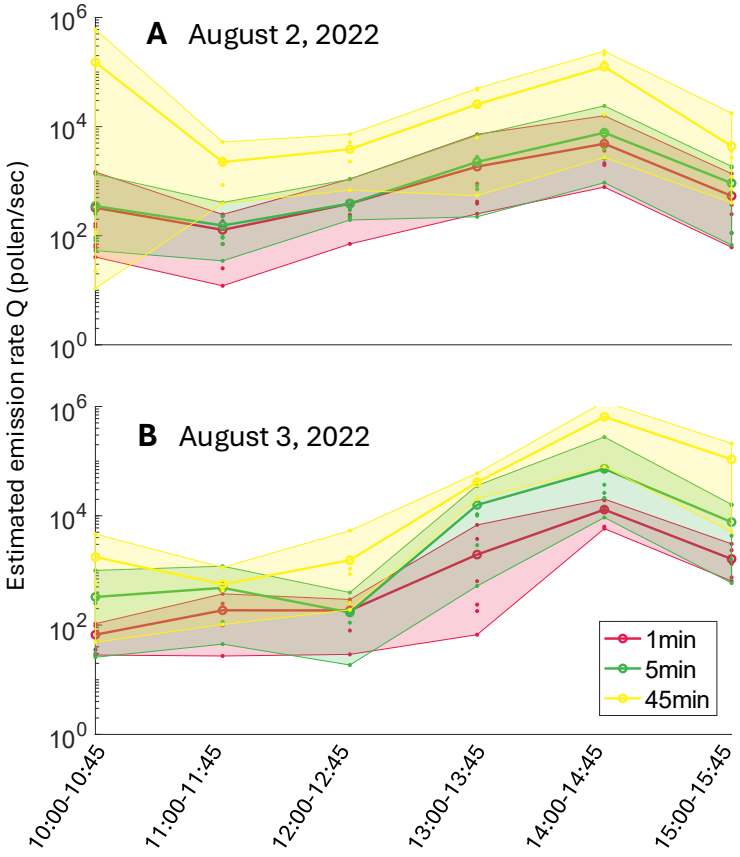


Figure 3.7: Mean estimated pollen emission rates for each sampling interval using 1-minute, 5-minute, and 45-minute simulations. (A) Estimated emission rates for August 2, 2022 and (B) August 3, 2022. Non-zero emission rate estimates for each sampler are shown as solid points. Shaded regions indicate the range the between minimum and maximum non-zero emission rate estimates. The vertical axis is on a log scale.

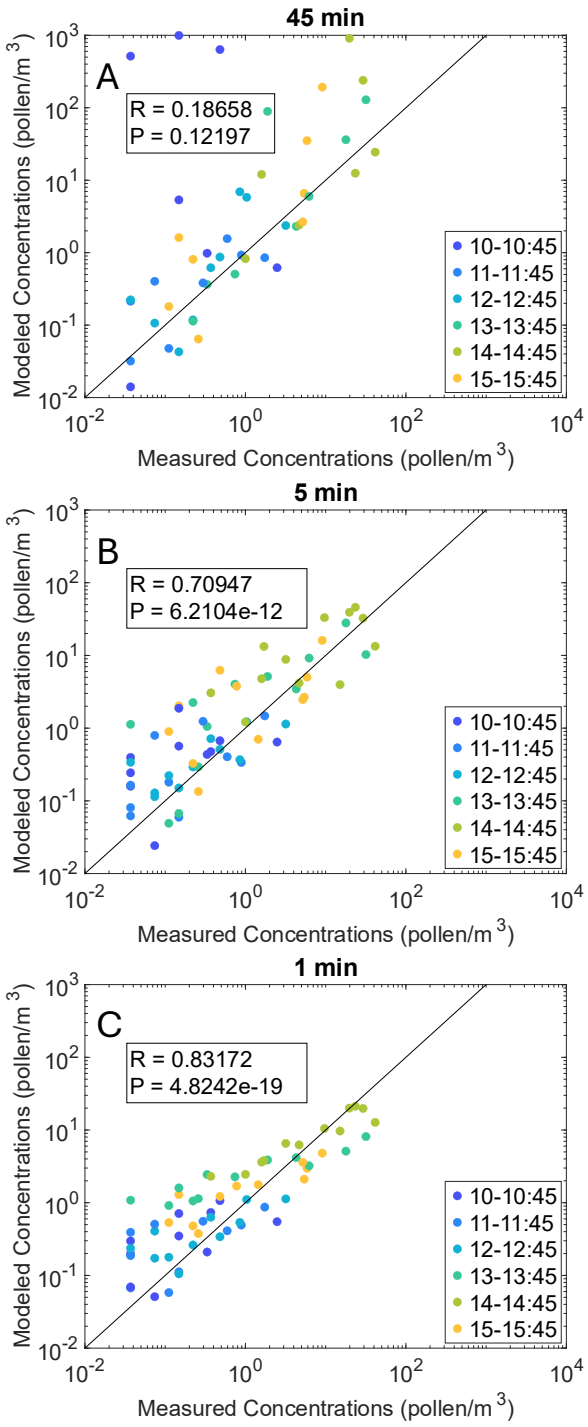


Figure 3.8: Measured concentrations at the ED samplers compared with simulated concentrations for (A) 45-minute, (B) 5 minute, and (C) one-minute intervals. These plots are not intended as model validation, but rather to show that decreasing the averaging time for simulations greatly improves modeled concentrations. Note that the plots are on log-log scales.

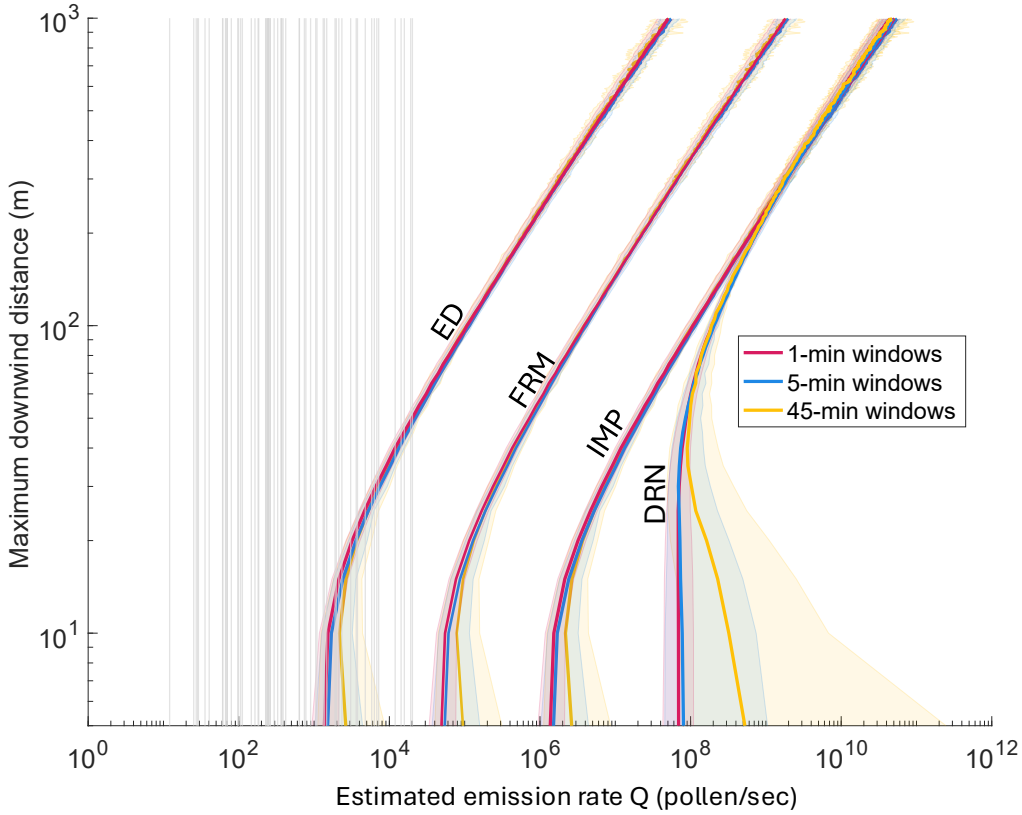


Figure 3.9: Maximum distance at which ED, FRM, IMP, and DRN sensors should be placed from the source to capture at least 100 pollen grains for a given a emission rate from the field. These are calculated using concentration curves derived from each sampling interval, distinguishing between estimates computed with 1-minute, 5-minute, and 45-minute averaging windows. Solid lines represent the median values, while the shaded regions indicate the range between the lowest and highest values observed across all sampling intervals. Gray vertical lines correspond to the estimated emission rates for each sensor during all sampling periods.

Chapter 4

Lagrangian Stochastic Model Evaluation in Convective Conditions

4.1 Introduction

Large-eddy simulation (LES) provides a high-fidelity time-varying description of turbulent transport and deposition, but its computational cost limits its routine use for applications that require many scenarios or long integration times. In contrast, Lagrangian stochastic (LS) dispersion models are inexpensive and widely used, yet their accuracy depends sensitively on flow regime and model assumptions. In this study, we quantitatively compare LES and two LS formulations across a range of boundary layer regimes to assess when LS can reasonably replace LES for predicting pollen and more general bioaerosol deposition patterns within 1 km of a ground-level source, a question of practical importance for gene flow caused by pollen and seed dispersal and the spread of plant disease via fungal spores [9, 27, 33, 35, 45].

The comparison in this study focuses on two established Lagrangian stochastic formulations that have been widely applied in atmospheric dispersion modeling. The first is mostly used for near-source dispersal, where large scale convective motions are not the main drivers of dispersal. It uses a jointly Gaussian distribution for the fluctuating velocity components (u', v', w'), which can then include the covariance $\overline{u'w'}$ between horizontal and vertical veloc-

ities [79, 88]. The assumption of gaussian vertical velocities is well-suited for shear-dominated turbulence at the surface, and has produced good agreement with near-source measurements of pollen and spore dispersal [29, 31, 53, 75, 89]. The second is a convective boundary-layer (CBL) formulation in which the vertical velocity is represented by a height-dependent, skewed bi-Gaussian distribution in order to better represent the updrafts and downdrafts of convective-turbulence-dominated boundary layers [86, 87, 119]. Both formulations were originally developed for one- or two-dimensional dispersal configurations and have been validated primarily against vertical profiles and along-wind concentration statistics rather than fully three-dimensional deposition distributions [31, 53, 54, 87, 100]. Conceptually, the SL model is most appropriate in the surface layer where turbulence is gaussian, produced mostly by shear, and the large coherent motions of convective turbulence are not present [31, 75, 79, 88], whereas the CBL model is designed for buoyancy-dominated, well-mixed convective boundary layers [86, 87, 119]. To isolate differences arising from these underlying stochastic formulations rather than from differing inputs, both LS models are driven here by the same set of analytical wind-statistics profiles across the range of stability regimes considered.

The LES simulations in this study are performed with the PALM large-eddy simulation model, which explicitly resolves the turbulent velocity field in the atmospheric boundary layer using a well-established framework [155, 156]. Particle dispersal is simulating using built-in Lagrangian Particle Dispersion Module (LPDM) included in the PALM package, in which particles are advected by the instantaneous resolved velocity field plus a stochastic subgrid-scale velocity increment following a Lagrangian stochastic formulation developed to be driven by the time-varying fields of large-eddy simulations [84, 155].

The PALM LES simulations conducted span a sequence of seven atmospheric boundary-layer regimes from near-neutral, high-wind, streak-dominated shear flows through roll-dominated mixed layers to deep, strongly convective boundary layers with large convective motions

[157, 158, 159]. The resulting time-averaged LES-driven deposition distributions are treated as the reference "truth" against which the SL and CBL LS formulations are evaluated.

To compare LES and LS performance, we focus on how well each model reproduces the LES plume shape for different deposition magnitudes. For each case, we compute the downwind and crosswind standard deviations of the ground-level deposition field, σ_x and σ_y, as primary measures of plume extent and lateral spread. We then assess the spatial fidelity of each LS formulation using a thresholded, binary comparison of the LES and LS deposition fields across a range of deposition levels. Spatial overlap at each threshold is quantified with the Jaccard index, defined as the ratio of intersection to union of two regions which exceed the threshold. Together, these metrics indicate whether the LS models reproduce the overall plume extent and lateral spread, and how much of the main deposition footprint they place in the same locations as the LES across different stability regimes.

In this paper, we first describe the LES configuration and the seven atmospheric boundary-layer regimes considered. We then summarize the SL and CBL LS formulations and the analytical wind-statistics profiles used to drive both of them. Next, we present a case-by-case comparison of LES and LS using deposition distributions, plume-spread metrics, and Jaccard overlap curves, highlighting where each model fails or succeeds. Finally, we discuss the implications for using LS as a substitute for LES in pollen and bioaerosol applications and outline key limitations and directions for future work.

4.2 Methodology

4.2.1 Representative ABL cases

To identify representative wind regimes, we used hourly meteorological output from a Weather Research and Forecasting (WRF) model simulation for year 2016 over the contiguous United States. This data was previously extracted for each US county from July to November in Nimmala et al. [123] and includes friction velocity u_*, Monin-Obukhov length L, convective velocity scale w_*, surface roughness length z_0, and boundary layer height z_i. The dataset was first filtered for $z_0 \approx 0.1$ m and negative Monin-Obukhov length. It was then partitioned into approximate Pasquill stability classes at $z_0 \approx 0.1$ m by $1/L$ following Golder (1972) [160]: neutral $(-0.01, 0]$, slightly unstable $(-0.035, -0.01]$, moderately unstable $(-0.095, -0.035]$, and strongly unstable ≤ -0.95. Within each stability class, cases were further grouped by u_* into low wind (0.1-0.2 $\mathrm{m\,s^{-1}}$), medium wind (0.3-0.4 $\mathrm{m\,s^{-1}}$), and high wind (> 0.5 $\mathrm{m\,s^{-1}}$). The remaining combinations were then binned by $w_* \approx 0.5, 1, 1.5$, and 2 $\mathrm{m\,s^{-1}}$, and the most frequently occurring boundary layer heights z_i were selected.

We estimated a kinematic heat flux using w_* and z_i [108], assuming a 300 K surface-potential temperature θ_0,

$$Q = \frac{{w_*}^3 \theta_0}{g z_i}. \tag{4.1}$$

These initial estimates were then adjusted toward values commonly used in convective ABL LES. For example, Pan et al. (2013) [161] used $Q = 0.025$ $\mathrm{K\,m\,s^{-1}}$ for weakly unstable cases, 0.05 for moderately unstable cases, and 0.1 for strongly unstable cases. In Weil et al. (2012) [85], the strongly convective case had $Q = 0.29$ $\mathrm{K\,m\,s^{-1}}$, while in Moeng and Sullivan (1994) [157] the buoyant case had $Q = 0.24$ $\mathrm{K\,m\,s^{-1}}$ and the two intermediate cases (strong

shear, moderate convection) had $Q = 0.05$ and $Q = 0.03$ K m s^{-1}. After adjusting Q for each regime, the corresponding z_i used in the LES was recomputed, and a geostrophic wind speed U_g was estimated using a logarithmic wind profile with the Paulson (1970) [162] stability function.

Above the boundary layer, the potential temperature profile was prescribed to increase by 1 K over 60 m for the neutral and weakly unstable cases and by 6 K over 60 m for the more strongly unstable cases. Above this capping inversion, all cases used a uniform lapse rate of 0.003 K m^{-1}. These values are consistent with previous convective LES studies [157, 163, 164].

Using this procedure, we define seven distinct daytime ABL simulations which fall into three stability regimes: shear-dominated near-neutral (cases 1 and 2), mixed shear and convective (cases 3, 4, and 5), and strongly convective highly unstable (cases 6 and 7). The initial parameters and final LES input values for these seven cases are summarized in Table 4.1.

Based on $-z_i/L$ values in Table 4.2, Case 1 lies in the precritical, near-neutral regime ($-z_i/L < 0.43$) identified by [158], in which turbulence is dominated by elongated low-speed streaks. Case 2 falls just beyond this threshold ($-z_i/L = O(1)$), where horizontal roll vortices begin to organize the flow as buoyancy becomes more comparable to shear [158]. Cases 3, 4, and 5 have $-z_i/L$ of order 10-30 and $w_*/u_* \approx$ 2-4, placing them firmly in convective but still shear-influenced mixed-layer regimes, consistent with the mixed shear-convective CBL structures described by [157, 159, 161, 165]; Case 3 is particularly similar to Moeng and Sullivans (1994) Case B, while Case 4 closely resembles the MUL02/MUL08 runs of [161]. Case 7 occupies a similar stability range but with a shallower, low-wind CBL and we have not found an exact analog in the literature. Finally, Cases 6 and 7 correspond to the most unstable conditions, with very large $-z_i/L$ and $w_*/u_* > 4$, similar to the deeply convective, buoyancy-dominated cases of [166], which are essentially free-convective regimes.

Case	Initial chosen ABL parameters			LES input parameters		
	$1/L$ (m^{-1})	u_* ($\mathrm{m\,s}^{-1}$)	w_* ($\mathrm{m\,s}^{-1}$)	z_i (m)	Q ($\mathrm{K\,m\,s}^{-1}$)	$U_{\mathrm{geostrophic}}$ ($\mathrm{m\,s}^{-1}$)
Case 1	-0.00061	0.60	0.5	380	0.01	13
Case 2	-0.00305	0.35	0.5	380	0.01	7
Case 3	-0.03391	0.30	1.0	440	0.07	4
Case 4	-0.02511	0.50	2.0	1030	0.24	8
Case 5	-0.38756	0.15	1.0	310	0.10	1.5
Case 6	-0.04844	0.30	1.5	1030	0.10	4
Case 7	-0.11627	0.30	2.0	1020	0.24	4

Table 4.1: Initial atmospheric boundary layer parameters and LES input parameters for the seven LES cases.

4.2.2 LES set-up

For the LES simulations, we use the PALM model, which is designed to simulate high-resolution atmospheric boundary layer flows and includes a built-in Lagrangian Particle Dispersion Model (LDPM) module [155, 156].

All cases are run on a 2km × 2km horizontal domain with a 2 km vertical extent, using $200 \times 200 \times 667$ grid points ($\Delta x = \Delta y = 10$ m, $\Delta z = 3$ m). Because we simulate a near-surface release and focus on ground-level concentrations and deposition, we adopt a relatively fine vertical grid to better resolve shear-dominated turbulence in the surface layer. Pan et al. [161] similarly used $\Delta z = 3$ m for a 1 m release height in convective boundary layer simulations.

Although PALM supports grid nesting and vertical grid stretching, the documentation notes that nesting requires further testing and that vertical stretching is not yet fully supported for the LDPM. The chosen 2km domain with uniform $\Delta z = 3$m is balance between resolving surface-layer eddies and maintaining computational feasibility.

Simulations are initialized with a constant geostrophic wind U_g specified in Table 4.1. We

apply Dirichlet (no-slip) boundary conditions for velocity at the surface, and Dirichlet conditions at the top, constraining the top boundary to the prescribed geostrophic wind. The initial potential temperature profile is defined using a reference potential temperature of 300 K up to the defined boundary-layer height and a positive lapse rate above, specified in Table 4.1 and in Section 4.2.1. Convection in all cases is driven by a constant surface kinematic heat flux Q (Table 4.1). Potential temperature boundary conditions are Neumann at the lower boundary (required when prescribing a constant surface heat flux) and 'initial gradient' at the top, which maintains the initial potential temperature gradient there. All lateral boundaries are cyclic.

All simulations use a latitude of 40° N, approximately the center of the continental United States, for PALM computes the Coriolis parameters internally. The resulting flows are influenced by the Coriolis affect and exhibit the turning of wind direction with height characteristic of Ekman boundary layers [108], particularly in the shear-dominated cases, which produces a leaning of the plume [167]. Although this does not represent an idealistic boundary layer and is not included in the LS models, we retain the Coriolis forcing as it strongly influences realistic development of turbulence [157, 168].All simulations were run for a total of five hours of simulated-time. Each required an average of 20 hours of wallclock time using 250 CPU cores on Virginia Techs TinkerCliffs cluster.

To represent dispersion over a rough, homogeneous surface, we use a roughness length $z_0 =$ 0.1 m for all cases. This value lies near the upper end of typical ranges reported for long grass, farmland, and crop canopies [108, 145]. Using PALMs LDPM, we simulate a near-surface point source by continuously releasing particles at a height of 2 m at the domain center. This near-surface release is intended to represent emission from canopy-height for grasses and other plants; elevated releases such as for trees interact differently with boundary layer turbulence and are outside the scope of this chapter. Particles are assigned a diameter of 30

μm based on a typical pollen (hemp, switchgrass) diameter of 30 μm[60, 111], and a density ratio of particles to air of 0.0012. Particles are released for all five hours of the simulation, at a constant rate of 1000 particles per 0.1 s to approximate a constant emission rate. All domain boundaries for particles are set to 'absorb', so that particles are removed from the simulation when they exit the domain.

4.2.3 LS set-up

In this chapter, we use two Lagrangian stochastic (LS) formulations: a convective boundary-layer (CBL) model and a surface-layer (SL) model. The full mathematical formulations, including the Langevin coefficients and wind-statistics profiles, are given in Appendix A and B respectively.

The SL model was formulated by assuming a Gaussian distribution for the turbulent velocities, which is reasonable in the small-scale shear-dominated turbulence of the surface layer. More specifically, it uses a jointly Gaussian distribution in which the vertical and downwind velocities are correlated, explicitly including the covariance $\overline{u'w'}$. In convective and unstable conditions, the SL model can capture the initial spread and rise of the plume near the source due to shear, but it cannot represent the large-scale updrafts and downdrafts in a highly convective boundary layer, for which a convective-layer formulation is better suited. Nevertheless, at short ranges and low heights above the ground, the SL model has shown good agreement with measurements across a variety of wind conditions [31]. In this chapter, we apply the SL model across the seven daytime ABL regimes defined in Section 4.2.1, from near-neutral, high-wind conditions to strongly convective mixed layers.

The CBL model was originally formulated for strongly convective conditions [86], where turbulence in the bulk of the boundary layer is driven primarily by buoyancy. It represents

the skewed vertical velocity distribution in the convective boundary layer with a bi-Gaussian pdf, allowing it to reproduce plume rise from the surface in thermals and subsequent descent from the boundary-layer top as rising air cools [86, 87]. [87] developed a closure scheme for skewed turbulence that reduces to Gaussian turbulence in the limit of zero skewness [87], and Boehm et al. (2008) introduced a merged parameterization that combines surface-layer and convective-boundary-layer wind statistics [54]. Luhar (2002) extended the formulation to three dimensions by treating the horizontal and vertical velocity components as independent, with Gaussian pdfs for the horizontal components and a bi-Gaussian pdf for the vertical component [119]. Importantly, the CBL model does not include Reynolds-stress covariance terms such as $\overline{u'w'}$ that are known to be important in the surface layer [54], and can underpredict concentrations near the source at ground level [54]. Here we use the CBL formulation across the same seven regimes to contrast with the SL model.

Both models assume a stationary, horizontally homogeneous ABL, so that wind statistics depend only on height z. Vertical profiles of mean wind, velocity variances and covariances, and turbulence dissipation rate are constructed from these using a variety of similarity and mixed-layer scaling relations, and can be found in detail in Appendix B. For both models, across all cases, we apply the same wind statistics profiles so that we can better understand model differences independent of the driving wind statistics. Of particular note, we use a single set of merged convective-boundary-layer profiles for $\sigma_w^2(z)$ and $\varepsilon(z)$ following Boehm et al. [54], together with a height-varying profile for $\overline{u'w'}(z)$ [79]. These profiles are constructed to bridge shear-dominated surface-layer statistics near the ground with mixed-layer statistics aloft, so that as stability varies across cases, the underlying velocity statistics profiles also adjust. In the original formulations, the SL model is typically driven by surface-layer similarity profiles that are tailored to handle strong shear near the surface, whereas the CBL model is usually paired with vertically uniform mixed-layer statistics, characteristic of a well-mixed

CBL. Here, by instead applying the same merged $\sigma_w^2(z)$, $\varepsilon(z)$, and $\overline{u'w'}(z)$ profiles to both the SL and CBL runs, we intentionally remove differences in the wind-statistics forcing so that any differences in results reflect the differences between the two LS formulations. However, as shown in the Results and Discussion, this results in serious deficiencies of the SL formulation and overly narrow CBL plumes in the near-neutral, shear-dominated regimes.

Each LS simulation is driven by four ABL parameters output from the corresponding LES simulation: the friction velocity u_*, Monin-Obukhov length L, convective velocity scale w_*, and boundary-layer height z_i, as well as the same surface roughness length $z_0 = 0.1$ m used in the LES. The first four values are the time-averaged values from hour 4-5 of the LES simulation, and are summarized in Table 4.2, along with the non-dimensional ratios $-z_i/L$ and w_*/u_* in the last two columns. These quantify the ratio between convective and shear generated turbulence, or stability.

For each case, the LS simulations are performed assuming the downwind direction x is aligned with the mean wind at the source height, so that the mean crosswind velocity satisfies $\overline{V} = 0$. The LS domain matches the LES 2 km $\times$ 2 km horizontal domain and extends vertically from the surface to z_i. For each case, 100,000 particles are released from a point source located at the domain center at a height of $z = 2$ m, removed from the simulation at the domain boundaries, and the simulation is conducted until all particles are removed. The ground is treated as a depositing boundary, with particles removed when they reach $z = 0$, while the upper and lateral boundaries are treated as absorbing. Particles are assigned a constant gravitational settling velocity $v_s = 0.027m/s$ based on a typical pollen (hemp, switchgrass) diameter of 30 μm[60, 111] and a density of water. As hemp pollen is nearly spherical [111], Stoke's law provides a good approximation of settling velocity [75, 122]. Each LS simulation required approximately 20 minutes of wall-clock time using a single HPC core on Virginia Techs TinkerCliffs cluster. This is 15,000 times less computationally expensive than the LES

simulations on the same system.

Case	LES output / LS input parameters				Non-dimensional ratios	
	$1/L$ (m^{-1})	u_* ($\mathrm{m\,s}^{-1}$)	w_* ($\mathrm{m\,s}^{-1}$)	z_i (m)	$-z_i/L$	w_*/u_*
Case 1	-0.00033	0.7348	0.4955	384	0.13	0.67
Case 2	-0.00208	0.3973	0.4941	384	0.80	1.24
Case 3	-0.06473	0.2414	0.9904	444	28.66	4.10
Case 4	-0.03089	0.4656	1.9981	1032	31.87	4.29
Case 5	-0.17917	0.1879	0.8889	312	55.93	4.73
Case 6	-0.08729	0.2460	1.4931	1032	90.18	6.07
Case 7	-0.16617	0.2646	1.9893	1023	170.33	7.52

Table 4.2: ABL parameters used to drive the Lagrangian stochastic (LS) models for the seven cases, sorted by non-dimensional stability ratios.

4.2.4 Comparison methods

In the LS model, we compute concentration as follows,

$$C(i,j,k) = Q_{flux}\frac{1}{V_{\mathrm{box}}}\frac{1}{N_{\mathrm{p}}}\sum_{n=1}^{N_{\mathrm{p}}} T_n(i,j,k), \tag{4.2}$$

where the sum time particles spend in each grid box $T_n(i,j,k)$ is normalized by the total number of particles released (N_{p}) and the volume of the grid box (V_{box} = 10 m × 10 m × 3 m), and then multiplied by the modeled emission rate (Q_{flux}). We output a "relative concentration" from the LS model by taking $Q_{flux} = 1$. Recall that this is a time-averaged relative concentration, as these models are stationary.

To output the same time-averaged relative concentration from the LES, we average the time-varying particle concentrations from hour 4 to 5 and divide it by the release rate (10000 particles/sec) and the grid volume (V_{box} = 10 m × 10 m × 3 m).

We then rotate this plume into the mean wind direction for each case, which is computed

by time-averaging the u and v velocity components from hour 4 to 5 of the LES simulation at the lowest grid-level (z = 3 m) at the point of particle release (x = y = 1000 m).

To compute deposition flux distribution (# particles deposited $m^{-2}s^{-1}$) from the ground-level concentrations (#particles $m^{-2}s^{-1}$ at the z = 3 m grid level) for both the LS and LES, we simply multiply this ground-level concentration by the settling velocity. For particles of this size, the deposition flux is almost entirely dependent on settling velocity [169].

Finally, in our analysis, we compute two metrics to quantify similarity between the plumes. The first is the standard deviation of deposition flux in the horizontal σ_x and vertical σ_y, a common measure of spread that is well-defined in [170]. We next compute the Jaccard Index over a number of thresholds, which allows us to compare the shape of the plume at different deposition flux magnitudes. For a given deposition flux threshold T, we construct binary maps for each model by assigning a value of 1 where the local deposition flux exceeds T and 0 otherwise. The Jaccard index at threshold T, $J(T)$, is defined as the ratio of the area where both binary maps equal 1 (intersection) to the area where at least one map equals 1 (union), so that $J(T) = 1$ indicates perfect spatial overlap at that threshold and $J(T) = 0$ indicates no overlap. Evaluating $J(T)$ over thresholds spanning several orders of magnitude yields curves that summarize how spatial agreement between LES and LS varies from the high-deposition plume core near the source to the low-deposition plume tail for each stability regime and LS formulation.

4.3 Results

In this section, we compare dispersion predicted by large-eddy simulation and Lagrangian stochastic models across the seven representative daytime ABL regimes introduced above. For each case, we first carry out a PALM-LES driven with surface heat flux Q, the geostrophic

wind speed U_g, and an initial potential temperature profile that defines the boundary-layer height z_i. We then then run two Lagrangian stochastic formulations—a surface-layer (SL) model and a convective boundary-layer (CBL) model using the LES-derived wind statistics summarized in Table 4.2—. The details of the PALM configuration, scalar release, and the SL and CBL Langevin equations are given in the preceding Methods sections. Here we focus on the resulting deposition fields and a set of simple, comparable summary statistics.

To organize the results, we group the cases by stability regime: near-neutral shear-dominated (cases 1 and 2), mixed shear and convective (Cases 3, 4, and 5), and strongly unstable (Cases 6 and 7). Within each regime, we use three measures to compare deposition flux (directly proportional to concentration in the lowest grid) between LES and LS. First, we compare heat maps of time-averaged deposition flux, which show plume shape, orientation, and the qualitative extent of downwind and crosswind transport. Next, we compute streamwise and crosswind plume spreads (σ_x, σ_y) from the ground-level concentration fields for LES and LS (Table 4.3. Lastly, we evaluate the Jaccard index between LES and LS as a function of concentration threshold, providing a scalar measure of spatial overlap over a range of plume cores and tails (Figure4.4). Together, these measures characterize plume location, spread, and spatial overlap.

4.3.1 Near-neutral cases

In the near-neutral regime cases 1 and 2 in Figure 4.1, the LES dispersal shows classic shear-dominated plumes. Depositions are elongated along the mean wind direction, with a long tail of low deposition and minimal cross-wind spread. These cases correspond to strong mean winds and weak buoyancy, so turbulence is generated primarily by shear rather than convective updrafts, and particles are transported far downwind before they can disperse

laterally. Cases 1 and 2 are qualitatively similar, but the very high wind case 1 produces a slightly narrower plume than the high-wind case 2. Both plumes have streamwise spreads of $\sigma_x \approx 200$ m, while the crosswind spread increases from $\sigma_y \approx 21$ m in Case 1 to $\sigma_y \approx 25$ m in case 2 (Table 4.3), consistent with a modest reduction in mean wind and slightly greater opportunity for lateral dispersion in case 2.

In contrast to LES, the surface-layer (SL) formulation shows practically no dispersal in either near-neutral case. For both cases 1 and 2, the SL deposition is confined to a very small region immediately downwind of the source. The resulting footprint is dramatically smaller than in the LES, indicating that in near-neutral conditions, the SL model as configured here, fails to produce a realistic plume.

The CBL formulation produces deposition plumes that extend a comparable distance downwind to the LES fields (Figure 4.1). The streamwise spread statistics in Table 4.3 reflect this: for case 1, the LES and CBL values of σ_x are 198.45 m and 236.38 m, and for case 2 they are 200.43 m and 232.80 m, corresponding to overestimates of about 19% and 16%. In the crosswind direction, however, the CBL formulation substantially underpredicts lateral spread. For case 1, the LES and CBL values of σ_y are 21.22 m and 5.03 m, and for case 2 they are 24.51 m and 6.67 m, corresponding to large underestimates of 73% and 76%. The SL formulation performs even worse in this regime, with $\sigma_x = 3.12$ m and 16.79 m and $\sigma_y = 3.01$ m and 4.00 m in cases 1 and 2, which are about 90–98% smaller than the LES values in the streamwise direction and 80–86% smaller in the crosswind direction. Overall, in near-neutral boundary layers, the CBL formulation reproduces the order of magnitude of streamwise spread but yields plumes that are too narrow laterally, whereas the SL formulation completely collapses both streamwise and crosswind spread for the near-neutral cases.

4.3.2 Mixed shear and convective cases

In the mixed shear and convective cases (3, 4, and 5) shown in Figure 4.2, buoyancy plays a more active role, and the LES plumes are less elongated and noticeably wider than in the near-neutral cases. In case 3 (Figure 4.2a), a moderately convective shallow boundary layer with medium wind results in less downwind transport and more lateral spread than in cases 1–2, with with $\sigma_x = 127.61$ m and $\sigma_y = 31.33$ m compared to $\sigma_x \approx 200$ m and $\sigma_y \approx 21 - 25$ m in the near-neutral regimes. In case 4 (Figure 4.2b), boundary layer is much deeper and convective velocity scale is larger, but the higher wind speed offsets these changes, yielding spreads $\sigma_x = 139.32$ m and $\sigma_y = 35.89$ m that are very similar to case 3. Cases 3 and 4 have similar shear-to-convective turbulence ratios, shown by the non-dimensional group $w_*/u_* = 4.1, 4.29$ respectively, which explains why their plumes are nearly identical. In case 5 (Figure 4.2c), the boundary layer is shallow and moderately convective, with the weakest wind speed of these three cases. The LES plume in case 5 shows comparable downwind spread ($\sigma_x = 145.40$ m) but the largest crosswind spread ($\sigma_y = 47.89$ m).

For the same moderately unstable and mixed cases (3, 4, and 5), both the SL and CBL LS formulations reproduce the main qualitative features of the LES plume: the plumes remain elongated downwind with noticeable crosswind spread, and their footprints extend over similar downwind distances as LES rather than collapsing around the source (Figure 4.2). Quantitatively, the downwind spread σ_x from the SL model is now sensible and much closer to LES. Across cases 3, 4, and 5, SL σ_x differs from the LES value by only about -10% to $+11\%$ (Table 4.3), showing that SL recovers the correct order of magnitude for downwind dispersion under moderately unstable conditions. The CBL formulation behaves similarly: it modestly over-predicts σ_x in case 4 by about 17%, and underpredicts it in cases 3 and 5 by roughly 13% and 20%, respectively. In contrast to the near-neutral regime, both LS

schemes therefore produce realistic downwind transport in moderately unstable conditions, with the SL formulation performing slightly better.

Crosswind spread remains systematically lower than LES for both LS formulations. In case 4, both models perform well: SL and CBL σ_y fall within about 8% and 6% of the LES value, respectively, so lateral spread is essentially captured. In cases 3 and 5, however, both schemes underpredict σ_y more strongly, and case 5 is the most challenging: SL underestimates LES σ_y by about 24%, while CBL underestimates it by about 35%. Overall, the SL and CBL formulations show comparable skill in lateral spread for case 4, with CBL slightly closer, whereas SL performs modestly better in cases 3 and 5. Consistent with Figure 4.2, the LS plumes in these moderately unstable and mixed cases are visibly broader than in cases 1–2 and show clear lateral dispersion about the centerline. Yet both formulations still tend to produce plumes that are too narrow compared with the wider, more irregular LES footprints, especially in case 5.

4.3.3 Strongly unstable cases

In the strongly unstable cases (6 and 7), buoyancy dominates over shear, and the LES plumes reflect classic convective boundary-layer behavior (Figure 4.3). A larger proportion of particles released near the surface are lofted upwards by convective updrafts, resulting in less downwind transport and shorter wider plumes compared to the moderately unstable cases (3, 4, and 5). Compared with the deep, shear-influenced convective case 4 and the shallow convective cases 3 and 5, cases 6 and 7 are both more strongly unstable and, especially for case 7. In case 6 (Figure 4.3a), a deep, strongly convective boundary layer with moderate mean wind still produces an elongated plume, but it is noticeably wider than previous cases, with $\sigma_x = 131.6$ m and $\sigma_y = 37.5$ m. In case 7 (Figure 4.3b), where the mean wind is the

same as in case 6, but the convective velocity is greater, convection plays and even greater role, yielding the largest crosswind spread of all seven cases ($\sigma_y = 52.6$ m) and the smallest downwind spread of $\sigma_x = 106.0$ m.

For these strongly unstable cases, both LS formulations strongly underpredict downwind spread σ_x relative to LES. In case 6, the SL and CBL models yield $\sigma_x = 96.3$ m and 93.1 m, corresponding to underestimates of about 27% and 29%, respectively. In case 7, the SL and CBL values are 91.1 m and 82.8 m, underpredicting LES by roughly 14% and 22%. Thus, while both schemes reproduce the correct order of magnitude for downwind dispersion in strongly unstable conditions, SL is slightly closer to LES than CBL for σ_x, and both models tend to shorten the plume considerably relative to LES.

Lateral spread is represented more accurately. In case 6, the SL and CBL crosswind spreads, $\sigma_y = 40.7$ m and 38.3 m, are within about +9% and +2% of the LES value, indicating that both formulations closely match the LES lateral dispersion, with CBL slightly closer. In case 7, SL gives $\sigma_y = 51.3$ m, only about 3% below LES, whereas CBL underestimates somewhat more strongly with $\sigma_y = 45.1$ m, about 14% lower. Compared with the near-neutral and moderately unstable regimes, the LS plumes in cases 6 and 7 therefore provide their best match to LES lateral spread: both formulations recover realistic crosswind plume widths, with SL performing slightly better overall across the two strongly unstable cases.

4.3.4 Comparing across models using the Jaccard Index

In Figure 4.4, we use the Jaccard index to quantify how well the LS models reproduce the two-dimensional deposition distribution from LES. For a given threshold, T, the Jaccard index $J(T)$ is defined as the ratio of the area where both LES and LS footprints exceed T to the area where at least one of them exceeds T. Thus, when $J(T) = 1$, there is perfect overlap

of the two plumes within that threshold T, and when $J(T) = 0$, there is no spatial overlap at all. We repeat this calculation at four thresholds (10^{-8}, 10^{-7}, 10^{-6}, and 10^{-5}), which are also annotated as contour lines in the deposition heat maps (Figures 4.1- 4.3). Figure 4.4 shows the Jaccard index $J(T)$ for each case at each threshold T for both LS formulations (SL is dotted, CBL is solid).

Figure 4.4 also includes box plots summarizing LES ensemble variability for case 3 specifically. We ran the case 3 LES five times with different random seeds and computed the same Jaccard index at each threshold between each of the ensembles and the initial simulation. These box plots represent the best attainable model performance under this setup by comparing LES to itself across ensemble members, and they serve as a benchmark to evaluate the LS models.

In Figure 4.4, the LES ensemble runs for case 3 have Jaccard indices ranging between 0.7-0.8, or a best possible overlap of only 70-80%. The 10^{-8} and 10^{-7} thresholds for case 3 are nearly on par with the ensemble values ($J(T) = 0.67$). The 10^{-8} threshold is the contour surrounding the outermost lowest-deposition parts of the distribution. When the Jaccard index for this threshold is high, it indicates that the LS captures the overall shape of the plume well. The higher thresholds (10^{-7}, 10^{-6}, 10^{-5}) successively delineate areas of the distribution closer to the center with more deposition. When these are high, it means that the center of the plume is also reproduced well. In Figure 4.4, low thresholds more frequently have high Jaccard indices, which shows that LS reproduces the overall shape of the deposition distribution more accurately than the center of the plume.

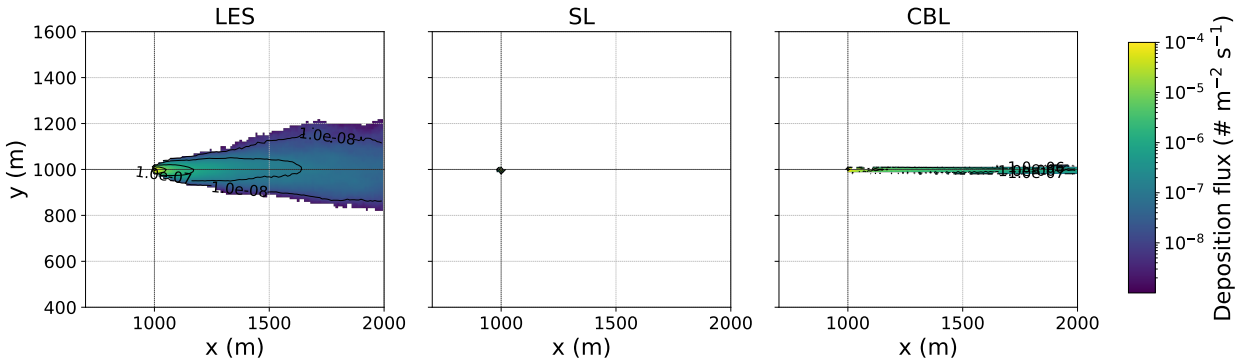
The Jaccard index trends across stability regimes mirror the spread statistics and heat maps. Cases 1 and 2—which completely collapse for the SL model and are overly narrow for the CBL model—have the lowest Jaccard indices, and the CBL model clearly performs better in Figure 4.4. We observe a general trend of increasing performance as cases become more

convection-dominated for the higher 10^{-7}, 10^{-6}, and 10^{-5} thresholds. In short, the more convective cases better capture the core of the deposition distribution, with some caveats. Namely, case 3 performs better than case 4, which has a dip in performance, followed by an increase for cases 5, 6, and 7. The performance dip for case 4 is not clear in the spread statistics, but the corresponding heat map shows that both SL and CBL models are over-predicting the downwind extent of deposition for these thresholds. This trend also does not hold for the highest 10^{-8}, which shows a steady decrease in Jaccard index for in cases 3–7. For the latter most convective cases 6 and 7, this is due to LS under-predicting the overall downwind extent of deposition.

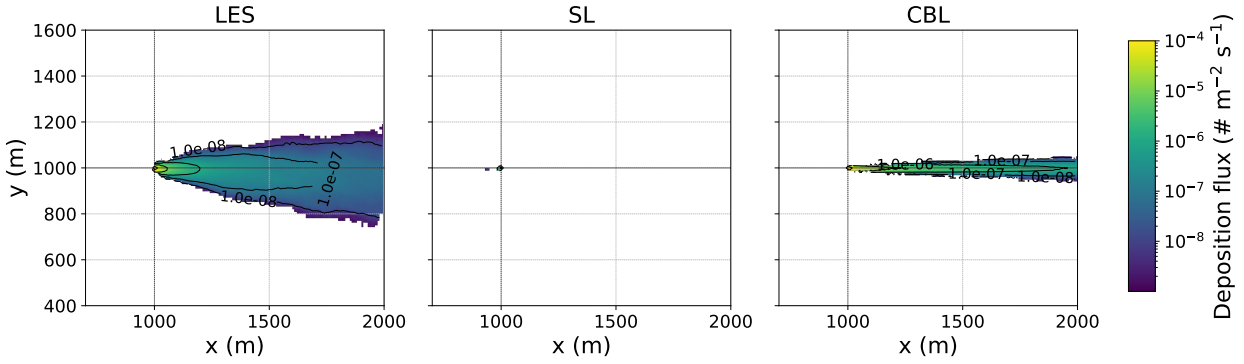
Finally, for the convective cases of interest, the two LS formulations are effectively interchangeable in terms of Jaccard performance. For Cases 3–6, the SL (dotted) and CBL (solid) curves almost lie on top of each other at each threshold, and the dominant variations in $J(T)$ come from the choice of threshold and stability regime rather than from the choice of LS formulation.

	LES		SL				CBL			
Case	σ_x	σ_y	σ_x	σ_y	$\%\Delta\sigma_x$	$\%\Delta\sigma_y$	σ_x	σ_y	$\%\Delta\sigma_x$	$\%\Delta\sigma_y$
1	198.4	21.2	3.1	3.0	-98.4	-85.8	236.4	5.0	19.1	-76.3
2	200.4	24.5	16.8	4.0	-91.6	-83.7	232.8	6.7	16.1	-72.8
3	127.6	31.3	114.8	25.4	-10.1	-19.1	110.8	24.3	-13.1	-22.3
4	139.3	35.9	153.9	33.0	10.5	-8.0	162.6	33.6	16.7	-6.4
5	145.4	47.9	134.6	36.3	-7.5	-24.1	115.7	31.2	-20.4	-34.8
6	131.6	37.5	96.3	40.7	-26.8	8.5	93.1	38.3	-29.2	2.1
7	106.0	52.6	91.1	51.3	-14.1	-2.5	82.8	45.1	-21.9	-14.3

Table 4.3: Downwind (σ_x) and crosswind (σ_y) spread statistics from LES and two Lagrangian stochastic (LS) formulations: surface-layer (SL) and convective boundary layer (CBL). Percentage differences $\Delta\sigma$ are relative to the LES spread.

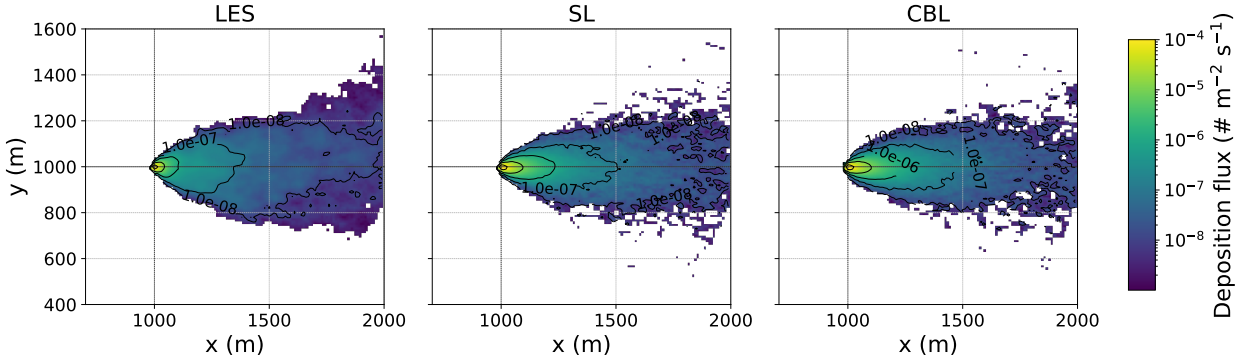


(a) Case 1

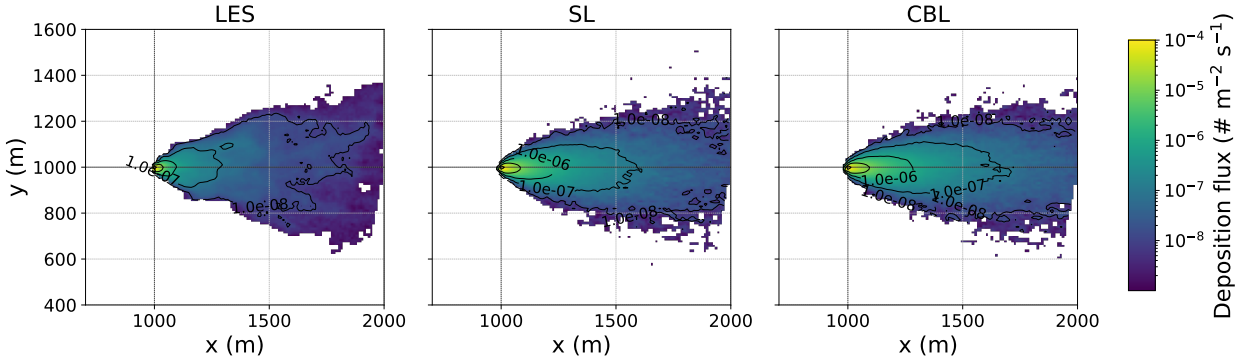


(b) Case 2

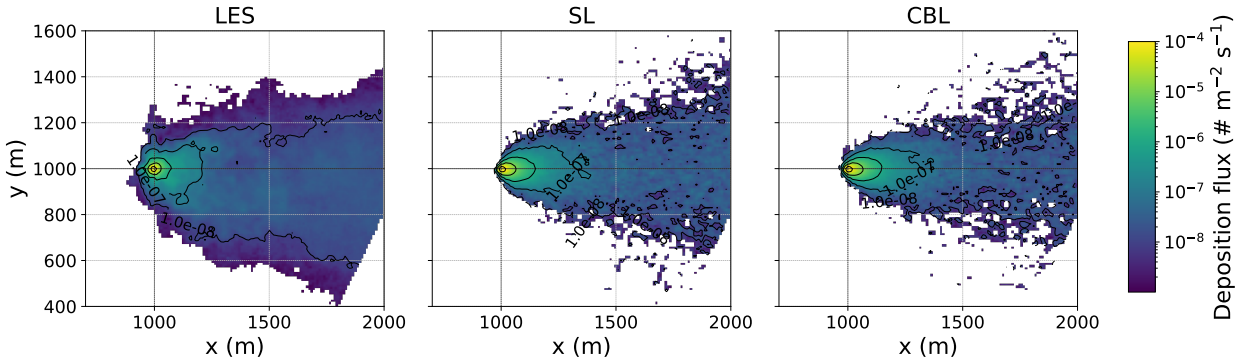
Figure 4.1: Deposition flux maps for the near-neutral cases (Cases 1–2) from LES, SL, and CBL models, rotated into the mean wind direction.



(a) Case 4

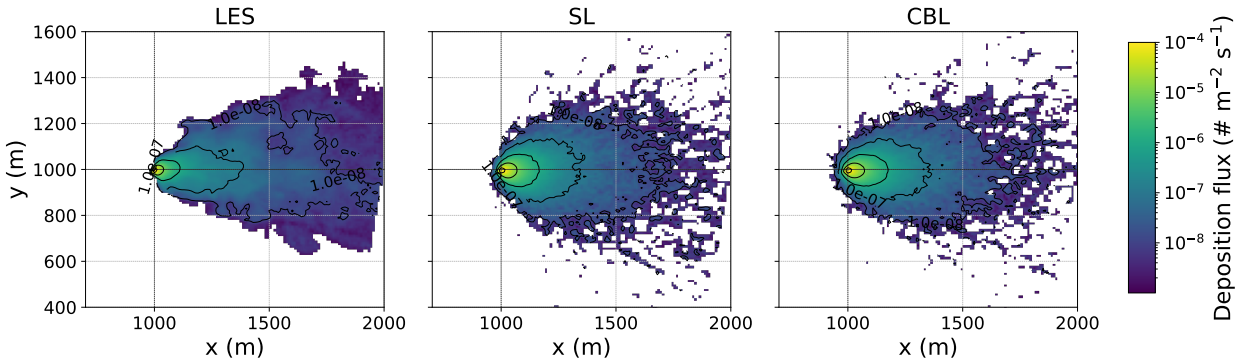


(b) Case 3

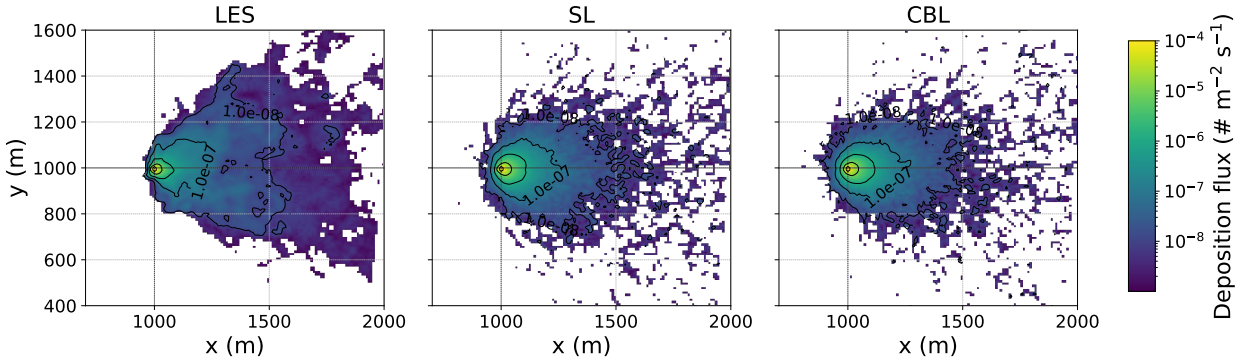


(c) Case 7

Figure 4.2: Deposition flux maps for moderately unstable to unstable cases (Cases 4, 3, and 7) from LES, SL, and CBL models, rotated into the mean wind direction.



(a) Case 5



(b) Case 6

Figure 4.3: Deposition flux maps for strongly unstable cases (Cases 5-6) from LES, SL, and CBL models, rotated into the mean wind direction.

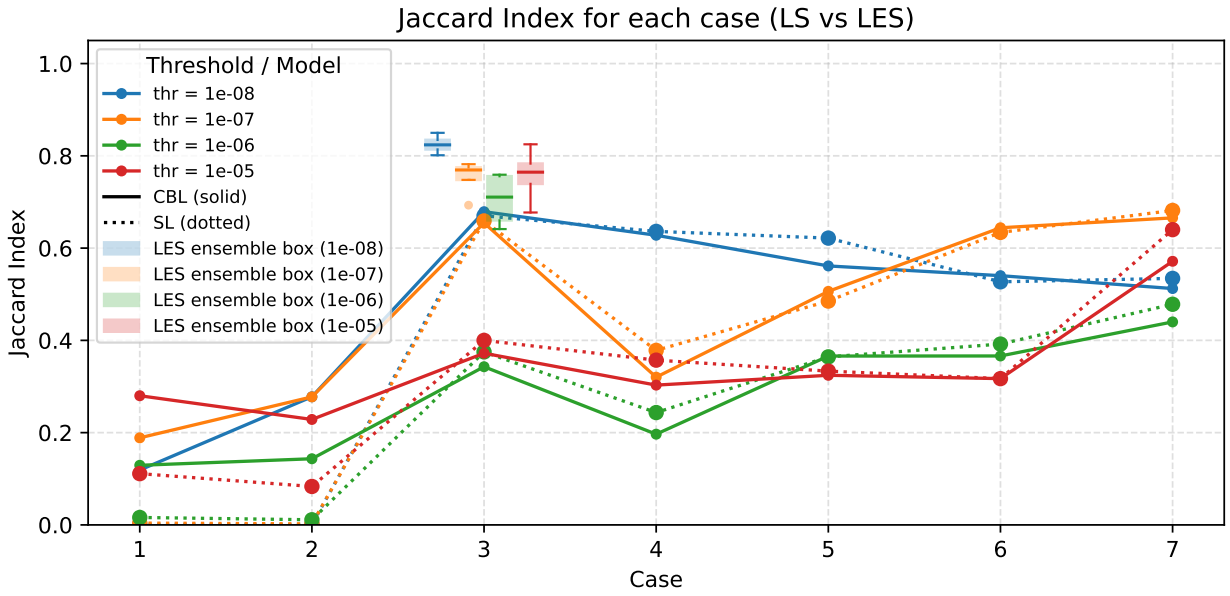


Figure 4.4: Jaccard index between the LES and the LS deposition distributions across Cases 1–7 for each concentration threshold. For a given threshold, the solid curve shows the convective boundary-layer (CBL) formulation and the dotted curve shows the surface-layer (SL) formulation (same color). Colored box plots show the variability of the LES ensembles for case 3 at each threshold, which demonstrates the best possible model performance for that case.

4.4 Discussion

All together, the seven LES simulations capture how the ground-level plume and associated deposition distributions evolve as the daytime boundary layer transitions from near-neutral, shear-dominated flow to deep, strongly convective ABL structures. In the near-neutral, high-wind Cases 1 and 2, the combination of small surface heat flux and $-z_i/L \ll 1$ means that turbulence in the surface layer is organized primarily into long, wind-aligned low-speed streaks, as in the neutral and weakly unstable shear-driven boundary layers of Moeng and Sullivan (1994) and Khanna and Brasseur (1998), and for Case 1, the subcritical ($-z_i/L < 0.43$) cases of Jayaraman and Brasseur (2020) [157, 158, 165, 166]. In Case 2, where $-z_i/L \approx 0.8$, buoyancy becomes slightly more comparable to shear, horizontal roll vortices appear [158], and the plume becomes somewhat wider than in Case 1 but still very elongated. In Cases 3 and 4, the boundary layer becomes definitively unstable (considered $-z_i/L > 20$) [165] as in the mostly convective regimes reported by Moeng and Sullivan (1994), Salesky et al. (2017), and Pan et al. (2013) [157, 159, 161]. In these cases, thermal updrafts break up the purely streak-dominated structure, and the ground-level plume becomes noticeably shorter and broader, with stronger cross-wind spread and faster vertical dilution than in Cases 1–2. Case 7 occupies a similar instability range but with a shallower, low-wind CBL which still allows substantial lateral spreading near the source. Cases 5 and 6 lie in strongly and very strongly convective, deep CBL regimes comparable to the most unstable simulations in Khanna and Brasseur (1998), and the near-free-convective LES of Salesky et al. (2017) [159, 166]. Here, the CBL structure completely transitions from rolls to cells, and the plume becomes very wide and relatively insensitive to the exact mean-wind speed. Salesky et al. (2017) [159] write that the most dramatic change in CBL structure happens up to $-z_i/L = 15 - 20$, which highlights a considerable gap in our cases. However, we still cover

several important regimes, and the changes in ground-level dispersion is notable.

Because all simulations use a near-surface release height of 2 m, the resulting plumes remain strongly coupled to surface-layer turbulence. While this is representative of a vegetative release, pollen released from trees is also often studied in allergenic and gene-flow contexts with release heights up to 16 m [44, 171, 172, 173], where pollen feels an increasing sensitivity to convective updrafts [85]. It is important to note that the model evaluation in this study is not generalizable to higher release heights. The comparison between models here applies only to near-surface particle emission.

In the near-neutral, shear-dominated Cases 1–2, the SL model fails because the jointly Gaussian velocity PDF that underlies its formulation [75, 79, 88] is no longer realizable with our prescribed statistics. This formulation assumes that u, v, and w are jointly Gaussian with variances and covariances defined by the Reynolds stress tensor, which must form a "symmetric, non-negative definite, and non-singular covariance matrix[79]". In the surface-layer formulation, this appears in, $A = 2\left(\sigma_u^2\sigma_w^2 - (\overline{u'w'})^2\right)$ which is only well-defined when $|\overline{u'w'}| < \sigma_u\sigma_w$. The merged surface-to-convective wind statistics profiles we use for $\sigma_w^2(z)$ and $\overline{u'w'}$ violate this constraint for the SL formulation, resulting in unrealistic vertical and horizontal particle velocities. Previous successful SL formulation applications did not have this issue [31, 53] as they used profiles for just the surface layer. We also make use of the SL formulation with solely surface-layer wind statistics in our field experiment (Chapter 4), and encounter no issues. Our results suggest that the SL framework is highly sensitive to wind statistics, and would require a sensitivity analysis to determine if it is actually inappropriate in shear-dominated ABL regimes.

Similarly, the CBL formulation faces issues in the shear-dominated cases 1 and 2 due to a different set of wind statistics. While it reproduces realistic downwind deposition distributions, it critically underestimates crosswind spread σ_y. For crosswind dispersion, both SL and CBL

formulations are entirely dependent on the variance of the crosswind velocity fluctuations, $\sigma_v^2(z)$, for which we use the Luhar (2002) convective scalings for the horizontal components, $\sigma_u^2 = \sigma_v^2 = (0.6w_*)^2$. This means that in both the CBL and SL runs the horizontal turbulent fluctuations are tied only to the convective velocity scale w_*, with no explicit dependence on the friction velocity u_* or shear-generated variances. In the near-neutral Cases 1–2, where u_* dominates and w_* is very small, this yields near-zero cross-wind velocity variance and both the CBL and SL formulations severely underpredict σ_y. In the definitively convective cases 3–7, where w_* is large and turbulence is well-mixed throughout the boundary layer, the same parameterization results in deposition distributions with crosswind spread that is much more consistent with the LES.

For downwind spread for cases 3–7, both formulations are consistent with the LES as well. Table 4.3 shows that, aside from the shear-dominated cases, the predicted downwind spreads from SL and CBL are very similar, with no systematic advantage of one formulation over the other. The SL and CBL models have considerably different formulations in the downwind direction; in the SL model, the downwind and vertical velocities are correlated, whereas in the CBL they are independent. In theory, this added complexity would mean that the SL might perform better in conditions where shear is a factor (cases 3-5). Our results show, however, that the dominant control on downwind spread in our setup is again the shared wind statistics rather than the detailed differences in the formulations.

Across all seven regimes, the Jaccard index clearly quantifies the skill with which SL and CBL formulations reproduce LES deposition distributions. It captures the collapse of the shear-dependent near-neutral cases 1 and 2, as well as performance trends for the more convective cases. In the more convective cases (3-7), the consequences of model choice are more nuanced. Both LS formulations reproduce the overall plume structure and footprint reasonably well: downwind and crosswind spreads are typically within 10-30% of the LES

values, and the Jaccard index curves in Figure 4.4 reach maxima of about 0.6-0.7 in the mixed and strongly convective regimes. In practical terms, that means that at certain deposition thresholds, roughly 60-70 % of the LES deposition area above that threshold is correctly identified by the LS model. For the more convective cases, both LS models reproduce the overall extent of the distribution more accurately than its inner structure, although this improves as cases become convection-dominated. For these cases, we also find that the two LS formulations perform similarly when comparing to the LES benchmark.

Atmospheric LES simulations are most-often used by researchers and modelers who are trying to better understand the physical phenomena of airborne dispersal. For operational users making real-time decisions, or for regulators and risk assessors who would require many simulations, LES is computationally prohibitive. Each of our LES simulations required approximately 20 hours of wallclock time using 250 CPU cores on Virginia Techs TinkerCliffs cluster, whereas each LS simulation finished within 20 minutes on a single core on the same cluster. This is 15,000 times less computationally expensive than the LES simulations and could easily be run on a personal computer.

The kinds of decisions that motivated this chapter (how far and wide a bioaerosol plume can spread near the ground) are typically made instead with lower-cost models. Applications include agricultural disease-forecasters [33], bioaerosols and particulates affecting human health [9, 35], and agricultural and environmental stakeholders looking to preempt, evaluate the risk of, or litigate gene flow [58, 91, 92].

If the wrong low-cost model is used in these settings, it might result in incorrect decision-making. For example, in near-neutral conditions, the LS formulations as presented in this chapter would yield overly narrow or collapsed plumes and thus underestimate the gene flow distances or disease spread. In strongly sheared, near-neutral boundary layers, neither formulation achieves sufficient overlap with LES to justify replacing a dedicated simulation,

without reevaluation of the wind statistics used to drive the models. In deep, strongly convective mixed layers, by contrast, a carefully calibrated LS model can reproduce the main qualitative features of the LES deposition distribution—overall extent, lateral footprint, and the location of the highest deposition—at a fraction of the computational cost. For decision-makers who care primarily about whether a receptor lies within a broad deposition zone (isolation distances around a pollen or disease source), this degree of agreement may be acceptable, especially when weighed against the orders-of-magnitude difference in computational cost.

In this context, the role of these LS models for pollen and bioaerosol deposition is best viewed as a controlled trade-off between computational requirements and accuracy. LS models are well-suited for rapid, low-cost mapping of the main plume envelope and for scenario screening in convective regimes. Because our Jaccard analysis is threshold-dependent, it also provides a more flexible way to use LS outputs in practice. A decision-maker can focus on the part of the plume and the range of deposition magnitudes that are operationally relevant and then read off how much of the LES benchmark the LS model captures at those thresholds. Given that the best Jaccard scores correspond to only moderate agreement, these LS models are most appropriate for tasks where relative patterns matter more than exact agreement, such as in early-stage risk assessment, sensitivity studies, or screening of management options. This study allows users to quantify and report the level of error in their LS simulations of deposition.

4.5 Limitations and Future Work

In this study, we deliberately used the same wind-statistics profiles to drive both the SL and CBL models in order to directly compare the formulations themselves. In theory, this

choice should make the comparison clearer, but in practice it produced unrealistic behavior in the SL model under strongly shear-dominated conditions. A more rigorous, deterministic comparison of the wind-statistics parameterizations—examining how different choices for mean profiles, variances, and covariances affect each model in the absence of stochastic noise—would help clarify the regimes in which each formulation remains mathematically and physically well behaved. It would also be useful to drive the LS models directly with time-averaged wind-statistics profiles directly from the LES, which would allow us to separate errors from the models themselves from those introduced by imperfect analytical wind-statistics parameterizations.

Beyond the wind statistics, several limitations remain. The SL and CBL schemes were originally developed for one- or two-dimensional dispersion and only later extended to three dimensions, so including an evaluation of vertical dispersion is needed, even though this chapter has focused on deposition patterns. Our seven LES cases sample near-neutral and strongly convective conditions but leave a gap across the transition from near-neutral to weakly and moderately unstable regimes, where both models begin to improve; additional cases in this range would better constrain where each LS formulation can credibly replace LES. Finally, the current comparison is restricted to a single ground-level point source and a downwind extent of order 1 km; extending the analysis to longer distances, elevated or distributed sources, and a broader range of particle properties would provide a more complete picture of when these 3D LS models can safely stand in for LES in bioaerosol and pollen-deposition applications.

4.6 Conclusion

This chapter has evaluated two Lagrangian stochastic (LS) particle dispersal models—a surface-layer (SL) formulation [31, 79, 88] and a convective boundary layer (CBL) formulation [86, 87, 119]—against LES-driven particle dispersal in seven different atmospheric boundary layer conditions spanning near-neutral, transitional, and strongly convective regimes. Specifically, we examined crosswind and downwind ground-level concentration, which is proportional to deposition flux. To the best of our knowledge, there is not a validation study for the 3-D extension of these LS models which also considers the lateral spread of the plume. We found that in the near-neutral, high-wind, shear-dominated cases, the particular merged wind-statistics parameterization used here caused the SL formulation to collapse the plume near the source and made the CBL plume unrealistically narrow, so that neither model provided an acceptable substitute for LES. In more strongly convective regimes, where the boundary layer is deep and buoyancy-dominated, both LS formulations produced plume shapes and deposition footprints that were qualitatively similar to LES, capturing the main extent and orientation of the deposition pattern when driven with wind statistics consistent with their underlying assumptions. Based on the metrics used here, the SL and CBL LS models achieve moderate but consistent agreement with LES. Still, their substantial computational efficiency means they remain highly useful for applications that require a full three-dimensional plume and deposition field but do not require perfect accuracy.

Chapter 5

Conclusions and future work

This dissertation integrates regional-scale Lagrangian stochastic (LS) modeling, local-scale field measurements, and large-eddy simulation (LES)-LS model comparisons to quantify and interpret windborne pollen transport and deposition across spatial scales relevant to cross-pollination, with broader applicability to any bioaerosol dispersal.

In Chapter 2, we characterized 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 used 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 revealed 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 found 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. 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.

The patterns we observed directly apply to any airborne particle of similar size ($\sim$ 30 μm), and can be used to identify regions and times of the year where those particles might travel farther and deposit in greater quantities. In a biosecurity and defense context, this is of vital interest to contain the spread of airborne agricultural diseases which can decimate fields, as well as human illnesses originating from animal or water bodies such as Legionnaires disease, Q-fever, and harmful algal blooms [9, 26, 27, 35]. These results could extend to any tracer released intentionally or not, to predict high-risk regions.

The Chapter 2 framework is intentionally simplified and leaves several open questions. More resolved, albeit more computationally expensive models, could incorporate detailed physics such as release of pollen from the anthers, dispersal within a canopy, wet deposition, and even conditions specific to a farms location like topography. Our vulnerability metric assumed one hemp pollen source per county. More detailed information on source locations and sizes would fuel a sophisticated vulnerability metric, or even an on-demand tool for risk-assessment for growers. Our study used meteorological data from 2016 only; performing this same study over multiple years would provide insight into yearly variation. Specifically, a long-term climatological study might reveal how warming temperatures and other trends change the seasonal spatial patterns we observed.

In Chapter 3, we moved from a regional-scale (up to 50 km) purely modeling approach to a highly local scale ($<$ 25 m) study that combines modeling with field measurements. In this study, three field campaigns were conducted around a very small field of genetically modified switchgrass tagged with orange-fluorescent protein for traceability. Switchgrass pollen was sampled up to 25 meters from the field using both traditional and novel volumetric samplers, including a drone-mounted sampler which collected pollen at a fixed height of 10 meters. Lagrangian stochastic dispersal simulations were run during sampling periods driven by high-

resolution wind measurements taken at the field. Our results from this unique experimental setup showcase several lessons that may guide future studies. Persistent low-wind meandering conditions were recorded throughout the campaign, and reducing the averaging window for simulations significantly improved pollen emission rate estimations by better incorporating shifting wind directions. The pollen emission rate was estimated by combining simulated and measured pollen concentrations, and strong diurnal trends were observed. Diurnal emission rate trends were positively correlated with wind speed, temperature, and vapor pressure deficit, and negatively correlated with relative humidity. This study highlights the potential for drone-based pollen sampling and GMO pollen tracking using fluorescence measurements. The findings provide insight into the effectiveness of different sensor types with respect to source strength and sampling distance, advancing the understanding of pollen dispersal dynamics and measurement techniques.

These results have important implications for allergen monitoring, cross-pollination risk assessment, and broader bioaerosol surveillance strategies. With unlimited resources, the natural next step would be a scaled up version of this same experiment focused on validation: a large, well-established source containing only PSYBIN1a to maximize the OFP signal, multi-year sampling on the same dates and times to quantify repeatability of diurnal emission patterns, and colocated Hirst-type gold-standard samplers deployed alongside the novel samplers to provide independent concentration estimates for model validation and rigorous sampler inter-comparison. A substantially larger source would allow meaningful comparisons between low- and high-volume samplers, with one set used to estimate the particle release rate and another for validating modeled concentrations. It would also enable more effective use of impinger-type samplers (IMP and DRN), which could preserve pollen for downstream viability studies, although isolated siting requirements for transgenic work would continue to pose limits for long-distance tracking. Focusing solely on PSYBIN1a switchgrass, with its

stronger OFP fluorescence in pollen, could further enhance tracking accuracy via automatic fluorescence-based quantification.

In Chapter 4, we evaluated two Lagrangian stochastic (LS) particle dispersal models, namely a surface-layer (SL) formulation [31, 79, 88] and a convective boundary layer (CBL) formulation [86, 87, 119], against an LES-driven stochastic particle dispersal model (PALM-LES) in seven different atmospheric boundary layer conditions spanning near-neutral, mixed shear and convective, and strongly convective regimes. To the best of our knowledge, there has not been a validation study for the 3-D extension of these LS formulations which also considers the lateral spread of the plume. In the near-neutral, shear-dominated cases, the merged wind-statistics parameterization caused the SL plume to collapse near the source and the CBL plume to become unrealistically narrow, so neither model was an acceptable substitute for LES. In more strongly convective, buoyancy-dominated regimes, both LS formulations produced plume shapes and deposition footprints that were qualitatively similar to LES and captured the main extent and orientation of the deposition pattern. Based on the metrics employed in this study, they achieved moderate agreement, so they are best viewed as computationally efficient tools for three-dimensional plume and deposition fields when perfect accuracy is not required.

In this study, we deliberately used the same wind-statistics profiles to drive both the SL and CBL models in order to directly compare the formulations themselves. In theory, this choice should make the comparison clearer, but in practice it produced unrealistic behavior in the SL model under strongly shear-dominated conditions. A more rigorous, deterministic comparison of the wind-statistics parameterizations—examining how different choices for mean profiles, variances, and covariances affect each model in the absence of stochastic noise—would help clarify the regimes in which each formulation remains mathematically and physically well behaved. It would also be useful to drive the LS models directly with

time-averaged wind-statistics profiles directly from the LES, which would allow us to separate errors from the models themselves from those introduced by imperfect analytical wind-statistics parameterizations. The current comparison is restricted to a single ground-level point source and a downwind extent of 1 km; extending the analysis to longer distances, elevated or distributed sources, and a broader range of particle properties would provide a more complete picture of when these 3D LS models can safely stand in for LES in bioaerosol and pollen-deposition applications.

These three chapters are organized around a single theme: using Lagrangian stochastic modeling, driven by realistic meteorology, to explore bioaersol dispersal at different scales—from regional dispersal patterns of cross-pollination risk, to what can actually be measured around a single small field, to when reduced-cost LS formulations can reasonably replace LES. In doing so, all three chapters operate at the intersection of engineering mechanics, atmospheric science, and plant biology: treating pollen as a mechanically transported particle in a turbulent boundary layer, while keeping the biological and agricultural context of hemp and switchgrass at the forefront. These results show how weather-informed modeling can support cross-pollination risk assessment and broader bioaerosol dispersal applications. It is a case-study for future collaborations across disciplines such as biosurveillance, agriculture, and human health.

Declaration of Generative AI

During the preparation of this work, the author used ChatGPT (OpenAI) to edit and improve original writing and to synthesize paragraphs from bullet points. it was used for feedback and brainstorming. Additioally, it was used for small coding tasks such as the creation of plots. After using this service, the author thoroughly reviewed and edited the content and takes full responsibility for the entirety of this work.

Bibliography

[1] Science First. 15000 High Volume Air Sampler. PDF product sheet, 2007. https://sciencefirst.com/wp-content/uploads/2017/05/15000-High-Volume-Air-Sampler.pdf.

[2] 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.

[3] VivianeR Després, J Alex Huffman, Susannah M Burrows, Corinna Hoose, AleksandrS Safatov, Galina Buryak, Janine Fröhlich-Nowoisky, Wolfgang Elbert, MeinratO Andreae, Ulrich Pöschl, et al. Primary biological aerosol particles in the atmosphere: a review. Tellus B: Chemical and Physical Meteorology, 64(1):15598, 2012.

[4] Kiran Kumari and Shweta Yadav. Unveiling the role of bioaerosols in climate processes: a mini review. International Journal of Environmental Research, 18(5):81, 2024.

[5] Manica Balant, Teresa Garnatje, Daniel Vitales, Marine Oganesian, Joan Vallès, Nina Stepanyan-Gandilyan, and Airy Gras. Bridging past and present: exploring cannabis traditions in armenia through ethnobotanical interviews and bibliographic prospecting. Journal of Cannabis Research, 7(1):8, 2025.

[6] Valentí Rull, Javier Sigro, and Teresa Vegas-Vilarrúbia. Cannabis pollen sources and dispersal in the iberian pyrenees during the last century: Preliminary results and proposals for future studies. Review of Palaeobotany and Palynology, 331:105208, 2024.

[7] Qingying Zhang, Xuan Chen, Hongyan Guo, Luisa M Trindade, Elma MJ Salentijn, Rong Guo, Mengbi Guo, Yanping Xu, and Ming Yang. Latitudinal adaptation and genetic insights into the origins of cannabis sativa l. Frontiers in plant science, 9: 420047, 2018.

[8] Moonil Kim, Seonghun Lee, Songhee Lee, Koong Yi, Hyung-Sub Kim, Sanghoon Chung, Junmo Chung, Hyun Seop Kim, and Tae Kyung Yoon. Seed dispersal models for natural regeneration: A review and prospects. Forests, 13(5):659, 2022.

[9] JPG Van Leuken, AN Swart, AH Havelaar, A Van Pul, W Van der Hoek, and D Heederik. Atmospheric dispersion modelling of bioaerosols that are pathogenic to humans and livestock–a review to inform risk assessment studies. Microbial Risk Analysis, 1: 19–39, 2016.

[10] Jessica N. Stockdale and Reginald J. Millwood. Transgene Bioconfinement: Dont Flow There. Plants, 12(5):1099, 2023. doi: 10.3390/plants12051099. URL https://doi.org/10.3390/plants12051099.

[11] Reginald Millwood, Madhugiri Nageswara-Rao, Rongjian Ye, Ellie Terry-Emert, Chelsea R Johnson, Micaha Hanson, Jason N Burris, Charles Kwit, and C Neal Stewart. Pollen-mediated gene flow from transgenic to non-transgenic switchgrass (Panicum virgatum L.) in the field. BMC Biotechnology, 17:1–10, 2017.

[12] John R Caradus. Is co-existence and/or containment of genetically modified plants possible, and is it important? New Zealand Journal of Agricultural Research, pages 1–52, 2025.

[13] Steven E Travers, D Bryan Bishop, and Cynthia L Sagers. Persistence of genetically engineered canola populations in the us and the adventitious presence of transgenes in the environment. Plos one, 19(5):e0295489, 2024.

[14] Francisco J Areal and Laura Riesgo. Eu inspections of gm content in food and feed: Are they effective? Agriculture, 11(9):842, 2021.

[15] Nicholas Kalaitzandonakes and Alexandre Magnier. The economics of adventitious presence thresholds in the eu seed market. Food Policy, 43:237–247, 2013.

[16] Jenny Teufel, Viviana López Hernández, Anita Greiter, Nele Kampffmeyer, Inga Hilbert, Michael Eckerstorfer, Frank Narendja, Andreas Heissenberger, and Samson Simon. Strategies for traceability to prevent unauthorised gmos (including ngts) in the eu: State of the art and possible alternative approaches. Foods, 13(3):369, 2024.

[17] C. Ahrens, G. Ecker, and C. Auer. The Intersection of Ecological Risk Assessment and Plant Communities: An Analysis of Agrostis and Panicum Species in the Northeastern U.S. Plant Ecol., 212(10):1629–1642, 2011. doi: 10.1007/s11258-011-9936-9.

[18] C. W. Ahrens, T. H. Meyer, and C. A. Auer. Distribution Models for Panicum virgatum (Poaceae) Reveal an Expanded Range in Present and Future Climate Regimes in the Northeastern United States. Am. J. Bot., 101(11):1886–1894, Nov 2014. doi: 10.3732/ajb.1400047. Epub 2014 Oct 24; PMID: 25366854.

[19] Charles Kwit and C Neal Stewart. Gene flow matters in switchgrass (Panicum virgatum L.), a potential widespread biofuel feedstock. Ecol. Appl., 22(1):3–7, 2012.

[20] Xiang Ren, Ting Cai, Zhongyuan Mi, Leonard Bielory, Christopher G Nolte, and Panos G Georgopoulos. Modeling past and future spatiotemporal distributions of airborne allergenic pollen across the contiguous united states. Frontiers in Allergy, 3: 959594, 2022.

[21] Yingxiao Zhang and Allison L Steiner. Projected climate-driven changes in pollen

emission season length and magnitude over the continental united states. Nature communications, 13(1):1234, 2022.

[22] Shubhayu Saha, Ambarish Vaidyanathan, Fiona Lo, Claudia Brown, and Jeremy J Hess. Short term physician visits and medication prescriptions for allergic disease associated with seasonal tree, grass, and weed pollen exposure across the united states. Environmental Health, 20(1):85, 2021.

[23] Mikhail Sofiev and Karl-Christian Bergmann. Allergenic pollen: a review of the production, release, distribution and health impacts. Springer Science & Business Media, 2012.

[24] Craig W Powers, Regina Hanlon, Hinrich Grothe, Aaron J Prussin, Linsey C Marr, and David G Schmale III. Coordinated sampling of microorganisms over freshwater and saltwater environments using an unmanned surface vehicle (USV) and a small unmanned aircraft system (sUAS). Front. Microbiol., 9:1668, 2018.

[25] Landon Bilyeu, Bryan Bloomfield, Regina Hanlon, Javier González-Rocha, Stephen J. Jacquemin, Andrew P. Ault, Johnna A. Birbeck, Judy A. Westrick, Hosein Foroutan, Shane D. Ross, Craig W. Powers, and David G. Schmale. Drone–Based Particle Monitoring above Two Harmful Algal Blooms (HABs) in the USA. Environ. Sci.: Atmos., 2(6):1351–1363, 2022. doi: 10.1039/D2EA00055E.

[26] Yung Sung Cheng, Yue Zhou, Richard H Pierce, Mike Henry, and Daniel G Baden. Characterization of florida red tide aerosol and the temporal profile of aerosol concentration. Toxicon, 55(5):922–929, 2010.

[27] Shane D Ross, Jeremie Fish, Klaus Moeltner, Erik M Bollt, Landon Bilyeu, and Tracy Fanara. Beach-level 24-hour forecasts of florida red tide-induced respiratory irritation. Harmful Algae, 111:102149, 2022.

[28] Donald E. Aylor. Biophysical scaling and the passive dispersal of fungus spores: relationship to integrated pest management strategies. Agricultural and Forest Meteorology, 97(4):275–292, 1999. ISSN 0168-1923. doi: https://doi.org/10.1016/S0168-1923(99)00072-6.

[29] Donald E Aylor, David G Schmale III, Elson J Shields, Maria Newcomb, and Carmen J Nappo. Tracking the potato late blight pathogen in the atmosphere using unmanned aerial vehicles and lagrangian modeling. Agricultural and Forest Meteorology, 151(2): 251–260, 2011.

[30] David Savage, Martin J Barbetti, William J MacLeod, Moin U Salam, and Michael Renton. Seasonal and diurnal patterns of spore release can significantly affect the proportion of spores expected to undergo long-distance dispersal. Microbial Ecology, 63:578–585, 2012.

[31] Donald E Aylor and Thomas K Flesch. Estimating spore release rates using a Lagrangian stochastic simulation model. Journal of Applied Meteorology and Climatology, 40(7):1196–1208, 2001.

[32] Marcelo Chamecki, Nicholas S Dufault, and Scott A Isard. Atmospheric dispersion of wheat rust spores: a new theoretical framework to interpret field data and estimate downwind dispersion. Journal of applied meteorology and climatology, 51(3):672–685, 2012.

[33] Hervé Van der Heyden, Pierre Dutilleul, Jean-Benoit Charron, Guillaume J Bilodeau, and Odile Carisse. Monitoring airborne inoculum for improved plant disease management. a review. Agronomy for Sustainable Development, 41(3):40, 2021.

[34] Meirui Qian and Jianli Jiang. Covid-19 and social distancing. Journal of Public Health, 30(1):259–261, 2022.

[35] Charles F Dillon and Michael B Dillon. Multiscale airborne infectious disease transmission. Applied and Environmental Microbiology, 87(4):e02314–20, 2021.

[36] Inland transport of aerosolized florida red tide toxins. Harmful Algae, 9(2):186–189, 2010. ISSN 1568-9883. doi: https://doi.org/10.1016/j.hal.2009.09.003.

[37] Abdul ZulAdly Mohaimin, Sarayu Krishnamoorthy, and Pooja Shivanand. A critical review on bioaerosolsdispersal of crop pathogenic microorganisms and their impact on crop yield. Brazilian Journal of Microbiology, 55(1):587–628, 2024.

[38] Donald E Aylor. The aerobiology of apple scab. Plant Disease, 82(8):838–849, 1998.

[39] James KM Brown and Mogens S Hovmøller. Aerial dispersal of pathogens on the global and continental scales and its impact on plant disease. Science, 297(5581): 537–541, 2002.

[40] O. Tackenberg, P. Poschlod, and S. Kahmen. Dandelion seed dispersal: The horizontal wind speed does not matter for long-distance dispersal - it is updraft! Plant Biology, 5(5):451454, 2003. ISSN 14358603. doi: 10.1055/s-2003-44789.

[41] Ran Nathan and Helene C. Muller-Landau. Spatial patterns of seed dispersal, their determinants and consequences for recruitment. Trends in Ecology Evolution, 15(7): 278–285, 2000. ISSN 0169-5347. doi: https://doi.org/10.1016/S0169-5347(00)01874-7.

[42] Michael L. Cain, Brook G. Milligan, and Allan E. Strand. Long-distance seed dispersal in plant populations. American Journal of Botany, 87(9):1217–1227, 2000. doi: https://doi.org/10.2307/2656714.

[43] James M Bullock and Ralph T Clarke. Long distance seed dispersal by wind: measuring and modelling the tail of the curve. Oecologia, 124(4):506–521, 2000.

[44] Anna Kuparinen, Tiina Markkanen, Hermanni Riikonen, and Timo Vesala. Modeling air-mediated dispersal of spores, pollen and seeds in forested areas. Ecological modelling, 208(2-4):177–188, 2007.

[45] Merel B. Soons, Gerrit W. Heil, Ran Nathan, and Gabriel G. Katul. Determinants of long-distance seed dispersal by wind in grasslands. Ecology, 85(11):30563068, 2004. ISSN 00129658. doi: 10.1890/03-0522.

[46] S. Joseph Wright, Ana Trakhtenbrot, Gil Bohrer, Matteo Detto, Gabriel G. Katul, Nir Horvitz, Helene C. Muller-Landau, Frank A. Jones, and Ran Nathan. Understanding strategies for seed dispersal by wind under contrasting atmospheric conditions. Proceedings of the National Academy of Sciences, 105(49):1908419089, Dec 2008. doi: 10.1073/pnas.0802697105. URL https://www.pnas.org/doi/full/10.1073/pnas.0802697105.

[47] Ran Nathan, Gabriel G Katul, Henry S Horn, Suvi M Thomas, Ram Oren, Roni Avissar, Stephen W Pacala, and Simon A Levin. Mechanisms of long-distance dispersal of seeds by wind. Nature, 418(6896):409–413, 2002.

[48] Ran Nathan and Gabriel G Katul. Foliage shedding in deciduous forests lifts up long-distance seed dispersal by wind. Proceedings of the National Academy of Sciences, 102 (23):8251–8256, 2005.

[49] Anna Kuparinen, Gabriel Katul, Ran Nathan, and Frank M. Schurr. Increases in air temperature can promote wind-driven dispersal and spread of plants. Proceedings of the Royal Society B: Biological Sciences, 276(1670):30813087, 2009. ISSN 14712970. doi: 10.1098/rspb.2009.0693.

[50] James M. Bullock, Steven M. White, Christel Prudhomme, Christine Tansey, Ramón Perea, and Danny A. P. Hooftman. Modelling spread of british wind-dispersed plants

under future wind speeds in a changing climate. Journal of Ecology, 100(1):104–115, 2012. doi: https://doi.org/10.1111/j.1365-2745.2011.01910.x.

[51] Nathalie Jarosz, Benjamin Loubet, Brigitte Durand, Xavier Foueillassar, and Laurent Huber. Variations in maize pollen emission and deposition in relation to microclimate. Environ. Sci. Technol., 39(12):4377–4384, 2005.

[52] Nathalie Jarosz, Benjamin Loubet, and Laurent Huber. Modelling airborne concentration and deposition rate of maize pollen. Atmospheric Environment, 38(33):5555–5566, 2004.

[53] Donald E Aylor, Matthew T Boehm, and Elson J Shields. Quantifying aerial concentrations of maize pollen in the atmospheric surface layer using remote-piloted airplanes and Lagrangian stochastic modeling. Journal of Applied Meteorology and Climatology, 45(7):1003–1015, 2006.

[54] Matthew T Boehm, Donald E Aylor, and Elson J Shields. Maize pollen dispersal under convective conditions. Journal of Applied Meteorology and Climatology, 47(1): 291–307, 2008.

[55] Alexis Marceau, Benjamin Loubet, Bruno Andrieu, Brigitte Durand, Xavier Foueillassar, and Laurent Huber. Modelling diurnal and seasonal patterns of maize pollen emission in relation to meteorological factors. Agric. For. Meteorol., 151(1):11–21, 2011.

[56] Hongwei Lu, Jingzhao Lu, and Li He. Modeling and estimation of pollen-mediated gene flow at the landscape scale. Ecological Indicators, 106:105500, 2019.

[57] Hugh J Beckie and Linda M Hall. Simple to complex: modelling crop pollen-mediated gene flow. Plant science, 175(5):615–628, 2008.

[58] María Luz Zapiola and Carol Ann Mallory-Smith. Pollen-mediated gene flow from transgenic perennial creeping bentgrass and hybridization at the landscape level. PLoS One, 12(3):e0173308, 2017.

[59] Lidia S Watrud, E Henry Lee, Anne Fairbrother, Connie Burdick, Jay R Reichman, Mike Bollman, Marjorie Storm, George King, and Peter K Van de Water. Evidence for landscape-level, pollen-mediated gene flow from genetically modified creeping bentgrass with cp4 epsps as a marker. Proceedings of the National Academy of Sciences, 101(40):14533–14538, 2004.

[60] Ernest Small and Tanya Antle. A preliminary study of pollen dispersal in Cannabis sativa in relation to wind direction. Journal of Industrial Hemp, 8(2):37–50, 2003.

[61] Baltasar Cabezudo, Marta Recio, JoséMaŕia Sánchez-Laulhé, María Del Mar Trigo, Francisco Javier Toro, and Fausto Polvorinos. Atmospheric transportation of marihuana pollen from North Africa to the southwest of Europe. Atmospheric Environment, 31(20):3323–3328, 1997.

[62] F Aznar, L Negral, S Moreno-Grau, B Elvira-Rendueles, I Costa-Gómez, and JM Moreno. Cannabis, an emerging aeroallergen in southeastern Spain (Region of Murcia). Science of The Total Environment, 833:155156, 2022.

[63] Ariel F Stein, Roland R Draxler, Glenn D Rolph, Barbara JB Stunder, Mark D Cohen, and Fong Ngan. Noaas hysplit atmospheric transport and dispersion modeling system. Bulletin of the American Meteorological Society, 96(12):2059–2077, 2015.

[64] K Wyat Appel, Sergey L Napelenok, Kristen M Foley, Havala OT Pye, Christian Hogrefe, Deborah J Luecken, Jesse O Bash, Shawn J Roselle, Jonathan E Pleim, Hosein Foroutan, et al. Description and evaluation of the community multiscale air

quality (cmaq) modeling system version 5.1. Geoscientific model development, 10(4): 1703–1732, 2017.

[65] Ignacio Pisso, Espen Sollum, Henrik Grythe, Nina I Kristiansen, Massimo Cassiani, Sabine Eckhardt, Delia Arnold, Don Morton, Rona L Thompson, Christine D Groot Zwaaftink, et al. The lagrangian particle dispersion model flexpart version 10.4. Geoscientific Model Development, 12(12):4955–4997, 2019.

[66] William C Skamarock, Joseph B Klemp, Jimy Dudhia, David O Gill, Zhiquan Liu, Judith Berner, Wei Wang, Jordan G Powers, Michael G Duda, Dale M Barker, et al. A description of the advanced research wrf version 4. NCAR tech. note ncar/tn-556+ str, 145(10.5065), 2019.

[67] Georg A Grell, Jimy Dudhia, David R Stauffer, et al. A description of the fifth-generation penn state/ncar mesoscale model (mm5). 1994.

[68] NOAA U.S. Department of Commerce National Centers for Environmental Prediction (NCEP), National Weather Service (NWS). Ncep north american mesoscale (nam) 12 km analysis. 2015. doi: 10.5065/G4RC-1N91. URL https://doi.org/10.5065/G4RC-1N91.

[69] Hans Hersbach, Bill Bell, Paul Berrisford, Shoji Hirahara, András Horányi, Joaquín Muñoz-Sabater, Julien Nicolas, Carole Peubey, Raluca Radu, Dinand Schepers, et al. The era5 global reanalysis. Quarterly journal of the royal meteorological society, 146 (730):1999–2049, 2020.

[70] Ronald Gelaro, Will McCarty, Max J Suárez, Ricardo Todling, Andrea Molod, Lawrence Takacs, Cynthia A Randles, Anton Darmenov, Michael G Bosilovich, Rolf Reichle, et al. The modern-era retrospective analysis for research and applications, version 2 (merra-2). Journal of climate, 30(14):5419–5454, 2017.

[71] Fedor Mesinger, Geoff DiMego, Eugenia Kalnay, Kenneth Mitchell, Perry C Shafran, Wesley Ebisuzaki, Dušan Jović, Jack Woollen, Eric Rogers, Ernesto H Berbery, et al. North american regional reanalysis. Bulletin of the American Meteorological Society, 87(3):343–360, 2006.

[72] Alex De Visscher. Air dispersion modeling: foundations and applications. John Wiley & Sons, 2013.

[73] Ying Pan, Marcelo Chamecki, and Scott A Isard. Large-eddy simulation of turbulence and particle dispersion inside the canopy roughness sublayer. Journal of fluid mechanics, 753:499–534, 2014.

[74] Yaxing Du, Christina Isaxon, Pontus Roldin, Kristoffer Mattisson, Sasu Karttunen, Xiaoyu Li, Ebba Malmqvist, and Leena Järvi. Large-eddy simulation of aerosol concentrations in a realistic urban environment: Model validation and transport mechanism. Environmental Pollution, 358:124475, 2024.

[75] D.E. Aylor. Aerial Dispersal of Pollen and Spores. American Phytopathological Society, 2017. ISBN 9780890545423. URL https://books.google.com/books?id=Z4e6swEACAAJ.

[76] Alan J Cimorelli, Steven G Perry, Akula Venkatram, Jeffrey C Weil, Robert J Paine, Robert B Wilson, Russell F Lee, Warren D Peters, and Roger W Brode. AERMOD: A dispersion model for industrial source applications. Part I: General model formulation and boundary layer characterization. Journal of Applied Meteorology and Climatology, 44(5):682–693, 2005.

[77] Andreas Stohl, C Forster, A Frank, P Seibert, and G Wotawa. The lagrangian particle dispersion model flexpart version 6.2. Atmospheric Chemistry and Physics, 5(9):2461–2474, 2005.

[78] Joseph S Scire, David G Strimaitis, Robert J Yamartino, et al. A users guide for the calpuff dispersion model. Earth Tech, Inc, 521:1–521, 2000.

[79] Howard C. Rodean. Stochastic Lagrangian Models of Turbulent Diffusion. Springer, 1996. ISBN 9781935704119.

[80] ACM Beljaars and AAM Holtslag. Flux parameterization over land surfaces for atmospheric models. Journal of Applied Meteorology and Climatology, 30(3):327–341, 1991.

[81] Frans TM Nieuwstadt. The turbulent structure of the stable, nocturnal boundary layer. Journal of Atmospheric Sciences, 41(14):2202–2216, 1984.

[82] Lakshmi H Kantha and Carol Anne Clayson. Small Scale Processes in Geophysical Fluid Flows. Elsevier, 2000.

[83] Geoffrey Ecker, Thomas Meyer, and Carol Auer. Pollen longevity and dispersion models for switchgrass. Crop Science, 53(3):11201127, 2013. ISSN 0011183X. doi: 10.2135/cropsci2012.06.0382.

[84] Jeffrey C Weil, Peter P Sullivan, and Chin-Hoh Moeng. The use of large-eddy simulations in lagrangian particle dispersion models. Journal of the atmospheric sciences, 61(23):2877–2887, 2004.

[85] Jeffrey C Weil, Peter P Sullivan, Edward G Patton, and Chin-Hoh Moeng. Statistical variability of dispersion in the convective boundary layer: Ensembles of simulations and observations. Boundary-layer meteorology, 145(1):185–210, 2012.

[86] Ashok K Luhar and Rex E Britter. A random walk model for dispersion in inhomogeneous turbulence in a convective boundary layer. Atmospheric Environment, 23(9): 1911–1924, 1989.

[87] Ashok K Luhar, Mark F Hibberd, and Peter J Hurley. Comparison of closure schemes used to specify the velocity PDF in Lagrangian stochastic dispersion models for convective conditions. Atmospheric Environment, 30(9):1407–1418, 1996.

[88] D. J. Thomson. Criteria for the selection of stochastic models of particle trajectories in turbulent flows. Journal of Fluid Mechanics, 180(2):529556, 1987. ISSN 14697645. doi: 10.1017/S0022112087001940.

[89] D. E. Aylor, N. P. Schultes, and E. J. Shields. An Aerobiological Framework for Assessing Cross-Pollination in Maize. Agric. For. Meteorol., 119(3-4):111–129, 2003. doi: 10.1016/S0168-1923(03)00159-X.

[90] Trey Malone and Kevin Gomez. Hemp in the United States: a case study of regulatory path dependence. Applied Economic Perspectives and Policy, 41(2):199–214, 2019.

[91] Colorado Department of Revenues Marijuana Enforcement Division and Colorado Department of Agriculture. HB21-1301 Cross Pollination Working Group Technical Report. Technical report, 2022. https://sbg.colorado.gov/med/1301-Work-Groups.

[92] Tanner McCarty and Jeffrey Young. Hemp production network effects: Are producers tipped toward suboptimal varietal selection by their neighbors? Journal of Applied Farm Economics, 3(2):4, 2020.

[93] Zoe Habekost. Defining Damage: The "Damage to Commercial Agricultural Products" Exception to Oregon's Right-to-Farm Law. Or. L. Rev., 100:181, 2021.

[94] Knut Faegri, Peter Emil Kaland, Knut Krzywinski, et al. Textbook of Pollen Analysis. Number Ed. 4. John Wiley & Sons Ltd., 1989.

[95] Canadian Seed Growers Association. Canadian regulations and procedures

for pedigreed seed crop production. https://seedgrowers.ca/seed-growers/requirements/, 2023.

[96] California Crop Improvement Association. Hemp Crop Standards (Cannabis sativa). https://ccia.ucdavis.edu/standards/crop-standards/hemp-crop-standards, 2019.

[97] Oregon CBD. Feminized Seed and the Ethics of Cannabis Farming. https://oregoncbdseeds.com/farmerresources/, 2017.

[98] Daniele Lagomarsino Oneto, Jacob Golan, Andrea Mazzino, Anne Pringle, and Agnese Seminara. Timing of fungal spore release dictates survival during atmospheric transport. Proceedings of the National Academy of Sciences, 117(10):51345143, 2020. ISSN 10916490. doi: 10.1073/pnas.1913752117.

[99] Anna Kuparinen. Mechanistic models for wind dispersal. Trends in Plant Science, 11 (6):296–301, 2006.

[100] Matthew T Boehm and Donald E Aylor. Lagrangian stochastic modeling of heavy particle transport in the convective boundary layer. Atmospheric Environment, 39 (27):4841–4850, 2005.

[101] US EPA. CMAQ Model Version 5.3 Input Data – 1/1/2016 - 12/31/2016 12 km CONUS, 2019.

[102] US EPA. Meteorological model performance for annual 2016 simulation WRF v3.8, 2019.

[103] Thomas Karl and Walter James Koss. Regional and national monthly, seasonal, and annual temperature weighted by area, 1895-1983, 1984.

[104] Beverley Adams-Groom, Carsten Ambelas Skjøth, Michael Baker, and Thomas E Welch. 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.

[105] Alexander Mendez and Mohammad Farazmand. Quantifying Rare Events in Spotting: How Far Do Wildfires Spread? Fire Safety Journal, 132:103630, Sep 2022. ISSN 0379-7112. doi: 10.1016/j.firesaf.2022.103630.

[106] GE Willis and JW Deardorff. A laboratory model of diffusion into the convective planetary boundary layer. Quarterly Journal of the Royal Meteorological Society, 102 (432):427–445, 1976.

[107] GA Briggs. Plume dispersion in the convective boundary layer. Part II: Analyses of CONDORS field experiment data. Journal of Applied Meteorology and Climatology, 32(8):1388–1425, 1993.

[108] Roland B Stull. An Introduction to Boundary Layer Meteorology. Springer, 1988.

[109] Namrta Choudhary, MB Siddiqui, Shazia Bi, and Sayyada Khatoon. Effect of seasonality and time after anthesis on the viability and longevity of Cannabis sativa pollen. Palynology, 38(2):235–241, 2014.

[110] Nadia Aboulaich, M Mar Trigo, Hassan Bouziane, Baltasar Cabezudo, Marta Recio, Mohamed El Kadiri, and Mohammed Ater. Variations and origin of the atmospheric pollen of Cannabis detected in the province of Tetouan (NW Morocco): 2008–2010. Science of the Total Environment, 443:413–419, 2013.

[111] Robert C Clarke. Botany of the Genus Cannabis. Haworth Press, Binghamton, NY, 1999.

[112] HA Borthwick and NJ Scully. Photoperiodic responses of hemp. Botanical Gazette, 116(1):14–29, 1954.

[113] Obed Quaicoe, Fafanyo Asiseh, and Omoanghe S Isikhuemhen. Qualitative analysis of industrial hemp production, markets, and sustainability in North Carolina, United States. Agriculture, 13(4):887, 2023.

[114] Timothy Coolong, Kate Cassity-Duffey, and Noelle Joy. Role of planting date on yield and cannabinoid content of day-neutral and photoperiod-sensitive hemp in Georgia, USA. HortTechnology, 33(1):138–145, 2023.

[115] Ravindra B Malabadi, Kiran P Kolkar, Raju K Chalannavar, L Lavanya, Gholamreza Abdi, et al. Cannabis sativa: Botany, cross pollination and plant breeding problems. International Journal of Research and Innovation in Applied Science, 8(4):174–190, 2023.

[116] Jeffrey Steiner. Making hemp a 21st century commodity in Oregonand beyond. Crops & Soils, 54(4):24–31, 2021.

[117] Jeffrey R Stokes, Rita Hartel, Linda B Ford, and Thomas B Casale. Cannabis (hemp) positive skin tests and respiratory symptoms. Annals of Allergy, Asthma & Immunology, 85(3):238–240, 2000.

[118] Jeffrey S Young and Tanner J McCarty. Adapting network theory for spatial network externalities in agriculture: A case study on hemp cross-pollination. American Journal of Agricultural Economics, 105(4):1267–1287, 2023.

[119] Ashok K. Luhar. The influence of vertical wind direction shear on dispersion in the convective boundary layer, and its incorporation in coastal fumigation models. Boundary-Layer Meteorology, 102(1):138, 2002. ISSN 00068314. doi: 10.1023/A:1012710118900.

[120] Ernest Small. Evolution and classification of Cannabis sativa (marijuana, hemp) in relation to human utilization. The Botanical Review, 81:189–294, 2015.

[121] Prabha Amarasinghe, Camille Pierre, Mahta Moussavi, Addisie Geremew, Selamawit Woldesenbet, and Aruna Weerasooriya. The morphological and anatomical variability of the stems of an industrial hemp collection and the properties of its fibres. Heliyon, 8(4), 2022.

[122] James S. Borrell. Rapid assessment protocol for pollen settling velocity: Implications for habitat fragmentation. Bioscience Horizons, 5:19, 2012. ISSN 17547431. doi: 10.1093/biohorizons/hzs002.

[123] Manu Nimmala, Shane D Ross, and Hosein Foroutan. Cannabis pollen dispersal across the United States. Sci. Rep., 14(1):20605, 2024.

[124] David J Parrish and John H Fike. The biology and agronomy of switchgrass for biofuels. Crit. Rev. Plant Sci., 24(5-6):423–459, 2005.

[125] Geoffrey Ecker, Juan Zalapa, and Carol Auer. Switchgrass (Panicum virgatum l.) genotypes differ between coastal sites and inland road corridors in the Northeastern US. PLoS One, 10(6):e0130414, 2015.

[126] Carol Auer, Thomas Meyer, and Vernie Sagun. Reducing Pollen Dispersal using Forest Windbreaks. Plant Science Articles, (28), 2016. URL https://digitalcommons.lib.uconn.edu/plsc_articles/28.

[127] Albert P Kausch, Joel Hague, Melvin Oliver, Yi Li, Henry Daniell, Peter Mascia, Lidia S Watrud, and C Neal Stewart. Transgenic perennial biofuel feedstocks and strategies for bioconfinement. Biofuels, 1(1):163–176, 2010.

[128] Scott A Isard and Stuart H Gage. Flow of Life in the Atmosphere. Michigan State University Press, 2001.

[129] Tim R Gottwald and W Louis Tedders. A spore and pollen trap for use on aerial remotely piloted vehicles. Phytopathology, 75(7):801–807, 1985.

[130] David G Schmale III, Benjamin R Dingus, and Charles Reinholtz. Development and application of an autonomous unmanned aerial vehicle for precise aerobiological sampling above agricultural fields. J. Field Robot., 25(3):133–147, 2008.

[131] Laszlo Techy, David G Schmale III, and Craig A Woolsey. Coordinated aerobiological sampling of a plant pathogen in the lower atmosphere using two autonomous unmanned aerial vehicles. J. Field Robot., 27(3):335–343, 2010.

[132] J. T. M. Buters, C. Antunes, A. Galveias, K.-C. Bergmann, M. Thibaudon, C. Galán, C. Schmidt-Weber, and J. Oteros. Pollen and spore monitoring in the world. Clin. Transl. Allergy, 8:9, Apr 2018. doi: 10.1186/s13601-018-0197-8.

[133] J. T. M. Buters, B. Clot, C. Galán, R. Gehrig, S. Gilge, F. Hentges, D. O'Connor, B. Sikoparija, C. Skjoth, and F. Tummon. Automatic Detection of Airborne Pollen: An Overview. Aerobiologia, 40(1):13–37, 2024. doi: 10.1007/s10453-022-09750-x.

[134] Chiara Suanno, Iris Aloisi, Delia Fernández-González, and Stefano Del Duca. Monitoring techniques for pollen allergy risk assessment. Environ. Res., 197:111109, 2021.

[135] Fiona Tummon, Lucas Alados Arboledas, Maira Bonini, Benjamin Guinot, Martin Hicke, Christophe Jacob, Vladimir Kendrovski, William McCairns, Eric Petermann, Vincent-Henri Peuch, et al. The need for Pan-European automatic pollen and fungal spore monitoring: A stakeholder workshop position paper. Clin. Transl. Allergy, 11 (3):e12015, 2021.

[136] Maria Pilar Plaza, Franziska Kolek, Vivien Leier-Wirtz, Jens Otto Brunner, Claudia Traidl-Hoffmann, and Athanasios Damialis. Detecting airborne pollen using an automatic, real-time monitoring system: Evidence from two sites. Int. J. Environ. Res. Public Health, 19(4):2471, 2022.

[137] S. Adamov, N. Lemonis, B. Clot, B. Crouzy, R. Gehrig, M.-J. Graber, C. Sallin, and F. Tummon. On the Measurement Uncertainty of Hirst-Type Volumetric Pollen and Spore Samplers. Aerobiologia, 40(1):77–91, 2024. doi: 10.1007/s10453-021-09724-5.

[138] Ruyu Li and Rongda Qu. High throughput agrobacterium-mediated switchgrass transformation. Biomass Bioenergy, 35(3):1046–1054, 2011.

[139] David G. J. Mann, Peter R. LaFayette, Laura L. Abercrombie, Zachary R. King, Mitra Mazarei, Mathew C. Halter, Charleson R. Poovaiah, Holly Baxter, Hui Shen, Richard A. Dixon, W. A. Parrott, and Jr. Stewart, C. Neal. Gateway-compatible vectors for high-throughput gene functional analysis in switchgrass (Panicum virgatum L.) and other monocot species. Plant Biotechnol. J., 10(2):226–236, Feb 2012. doi: 10.1111/j.1467-7652.2011.00658.x. Epub 2011 Sep 28; PMID: 21955653.

[140] J Hollis Rice, Reginald J Millwood, Richard E Mundell, Orlando D Chambers, Laura L Abercrombie, H Maelor Davies, and C Neal Stewart. An orange fluorescent protein tagging system for real-time pollen tracking. BMC Res. Notes, 6:1–7, 2013.

[141] ARA Instruments. N-FRM Sampler. https://arainstruments.com/products/n-frm-sensor/, 2016. Product web page.

[142] Powers, C. Schmale-Lab-3D-Printing-Files-Powers-et-al-Frontiers-2018. GitHub repository, 2018. https://github.com/SchmaleLab/Schmale-Lab-3D-Printing-Files-Powers-et-al-Frontiers-2018.

[143] P Schotanus, FTMf Nieuwstadt, and HAR De Bruin. Temperature measurement with a sonic anemometer and its application to heat and moisture fluxes. Boundary-Layer Meteorol., 26:81–93, 1983.

[144] KM Lin, JY Juang, Y-W Shiu, and LFW Chang. Estimating the Bowen ratio for application in air quality models by integrating a simplified analytical expression with measurement data. J. Appl. Meteorol. Climatol., 55(4):1041–1048, 2016.

[145] Frank V Hansen. Surface roughness lengths. US Army Research Laboratory, 1993.

[146] Thomas K Flesch, John D Wilson, and Eugene Yee. Backward-time Lagrangian stochastic dispersion models and their application to estimate gaseous emissions. J. Appl. Meteorol. Climatol., 34(6):1320–1332, 1995.

[147] B. Addison-Smith, D. Wraith, and J. M. Davies. Standardising Pollen Monitoring: Quantifying Confidence Intervals for Measurements of Airborne Pollen Concentration. Aerobiologia, 36(4):605–615, 2020. doi: 10.1007/s10453-020-09656-6.

[148] Binbin Lin, Amir Bozorgmagham, Shane D Ross, and David G Schmale III. Small fluctuations in the recovery of fusaria across consecutive sampling intervals with unmanned aircraft 100 m above ground level. Aerobiologia, 29:45–54, 2013.

[149] David J OConnor, David A Healy, Stig Hellebust, Jeroen TM Buters, and John R Sodeau. Using the WIBS-4 (Waveband Integrated Bioaerosol Sensor) technique for the on-line detection of pollen grains. Aerosol Sci. Technol., 48(4):341–349, 2014.

[150] Dean Vickers, Larry Mahrt, and Danijel Belušić. Particle simulations of dispersion using observed meandering and turbulence. Acta Geophys., 56:234–256, 2008.

[151] D. Anfossi, D. Öttl, G. Degrazia, and L. A. Goulart. An Analysis of Sonic Anemometer

Observations in Low Wind Speed Conditions. Boundary-Layer Meteorol., 114(1):179–203, 2005. doi: 10.1007/s10546-004-1984-4.

[152] D. Anfossi, G. Brusasca, and G. Tinarelli. Simulation of Atmospheric Diffusion in Low Windspeed Meandering Conditions by a Monte Carlo Dispersion Model. Nuovo Cim. C, 13(6):995–1006, 1990. doi: 10.1007/BF02514787.

[153] J. F. Sagendorf and C. R. Dickson. Diffusion under low wind-speed inversion conditions. Technical memorandum erl arl-52, U.S. National Oceanic and Atmospheric Administration, Air Resources Laboratory, Silver Spring, MD, 1974. URL https://repository.library.noaa.gov/view/noaa/31330.

[154] Marcelo Chamecki, Charles Meneveau, and Marc B Parlange. Large eddy simulation of pollen transport in the atmospheric boundary layer. J. Aerosol Sci., 40(3):241–255, 2009.

[155] Björn Maronga, Micha Gryschka, Rieke Heinze, Fabian Hoffmann, Farah Kanani-Sühring, Marius Keck, K Ketelsen, Marcus Oliver Letzel, Matthias Sühring, and Siegfried Raasch. The parallelized large-eddy simulation model (palm) version 4.0 for atmospheric and oceanic flows: model formulation, recent developments, and future perspectives. Geoscientific Model Development, 8(8):2515–2551, 2015.

[156] Björn Maronga, Sabine Banzhaf, Cornelia Burmeister, Thomas Esch, Renate Forkel, Dominik Fröhlich, Vladimir Fuka, Katrin Frieda Gehrke, Jan Geletič, Sebastian Giersch, et al. Overview of the palm model system 6.0. Geoscientific Model Development, 13(3):1335–1372, 2020.

[157] Chin-Hoh Moeng and Peter P Sullivan. A comparison of shear-and buoyancy-driven planetary boundary layer flows. Journal of Atmospheric Sciences, 51(7):999–1022, 1994.

[158] Balaji Jayaraman and James G Brasseur. Transition in atmospheric boundary layer turbulence structure from neutral to convective, and large-scale rolls. Journal of Fluid Mechanics, 913:A42, 2021.

[159] Scott T Salesky, Marcelo Chamecki, and Elie Bou-Zeid. On the nature of the transition between roll and cellular organization in the convective boundary layer. Boundary-layer meteorology, 163(1):41–68, 2017.

[160] Donald Golder. Relations among stability parameters in the surface layer. Boundary-Layer Meteorology, 3(1):47–58, 1972.

[161] Ying Pan, Marcelo Chamecki, and Scott A Isard. Dispersion of heavy particles emitted from area sources in the unstable atmospheric boundary layer. Boundary-layer meteorology, 146(2):235–256, 2013.

[162] Ca A Paulson. The mathematical representation of wind speed and temperature profiles in the unstable atmospheric surface layer. Journal of Applied Meteorology and Climatology, 9(6):857–861, 1970.

[163] Peter P Sullivan, Chin-Hoh Moeng, Bjorn Stevens, Donald H Lenschow, and Shane D Mayor. Structure of the entrainment zone capping the convective atmospheric boundary layer. Journal of the atmospheric sciences, 55(19):3042–3064, 1998.

[164] M Abkar and P Moin. Les of the convective boundary layer: A minimum-dissipation modeling approach. Ann Res Brief, Center for Turbulence Research, Stanford, CA, USA, 2016.

[165] Alessandro Dosio, Jordi Vilà-Guerau de Arellano, Albert AM Holtslag, and Peter JH Builtjes. Dispersion of a passive tracer in buoyancy-and shear-driven boundary layers. Journal of Applied Meteorology, 42(8):1116–1130, 2003.

[166] Samir Khanna and James G Brasseur. Three-dimensional buoyancy-and shear-induced local structure of the atmospheric boundary layer. Journal of the atmospheric sciences, 55(5):710–743, 1998.

[167] Julia N Paegle and SK Kao. Turbulent diffusion in the ekman boundary layer. Atmospheric Environment (1967), 11(2):95–100, 1977.

[168] Tammy M Weckwerth, Thomas W Horst, and James W Wilson. An observational study of the evolution of horizontal convective rolls. Monthly weather review, 127(9): 2160–2179, 1999.

[169] Delphine K Farmer, Erin K Boedicker, and Holly M DeBolt. Dry deposition of atmospheric aerosols: Approaches, observations, and mechanisms. Annual review of physical chemistry, 72(1):375–397, 2021.

[170] D Finn, KL Clawson, RM Eckman, H Liu, ES Russell, Z Gao, and S Brooks. Project sagebrush: Revisiting the value of the horizontal plume spread parameter σ y. Journal of Applied Meteorology and Climatology, 55(6):1305–1322, 2016.

[171] Anna-Katharina Eisen, Barbara Fussi, Branko Šikoparija, and Susanne Jochner-Oette. Aerobiological pollen deposition and transport of fraxinus excelsior l. at a small spatial scale. Forests, 13(3):424, 2022.

[172] Yong Zhang, Leonard Bielory, and Panos G Georgopoulos. Climate change effect on betula (birch) and quercus (oak) pollen seasons in the united states. International journal of biometeorology, 58(5):909–919, 2014.

[173] Claire G Williams. Long-distance pine pollen still germinates after meso-scale dispersal. American Journal of Botany, 97(5):846–855, 2010.

[174] Michael Optis, Adam Monahan, and Fred C Bosveld. Limitations and breakdown of monin–obukhov similarity theory for wind profile extrapolation under stable stratification. Wind Energy, 19(6):1053–1072, 2016.

[175] RA Brost and JC Wyngaard. A model study of the stably stratified planetary boundary layer. Journal of Atmospheric Sciences, 35(8):1427–1440, 1978.

Appendices

Appendix A

Lagrangian Stochastic (LS) Model Formulations

A.1 Overview

The Lagrangian stochastic (LS) model is an application of Brownian motion to turbulent diffusion [79], 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.

The position increments for particles in the x (downwind) and z (vertical) directions are as follows [75],

$$dx = (u' + \overline{U})dt, \tag{A.1}$$

$$dz = (v' + \overline{V})dt, \tag{A.2}$$

$$dz = (w' - v_s)dt, \tag{A.3}$$

where u',v', and w' represent the fluctuating horizontal and vertical velocities, $\overline{U}$ is the mean horizontal wind velocity described further in Section B.3, and v_s is a constant settling velocity

for the particle.

Particle velocity increments[31] in the x and z directions are computed using the Langevin equation,

$$du' = a_u dt + b_u \mathcal{N}(0, dt), \tag{A.4}$$

$$dv' = a_v dt + b_v \mathcal{N}(0, dt), \tag{A.5}$$

$$dw' = a_w dt + b_w \mathcal{N}(0, dt), \tag{A.6}$$

which describes the incremental changes in u', v', and w' fluctuating particle velocities. The Langevin coefficients, a_u, a_v, a_w and b_u, b_v, 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[31, 87] of the lagrangian timescale τ:

$$dt = 0.02\tau, \tag{A.7}$$

$$\tau = 2\frac{\sigma_w^2}{C_0 \varepsilon}, \tag{A.8}$$

where we chose the constant $C_0 = 3$ [54, 119] and σ_w^2 is the vertical velocity variance and ε is the turbulent dissipation rate, described in Appendix B, which explicitly lists the wind statistics used in the LS simulations in each chapter.

In this dissertation, we often refer to two specific formulations of the LS model, which are described further in appendices A.3 and A.2. Both LS model formulations referenced in this dissertation are based on the well-mixed condition of Thomson (1987) [79, 86, 88], which requires "particles that are well-mixed remain so" [86, 88]. This means that the Eulerian velocity pdf $P_E(u_i', z)$ of the fluid, which is the probability of the fluctuating fluid velocity

components u', v', and w' taking on certain values as a function of position, should be exactly the same as the probability of fluctuating particle velocities u', v', and w' taking on the same values at the same positions, $P_a(u_i', z)$ [86]. With these assumptions, Thomson [88] derived the following general form for the Langevin coefficients from the Fokker-Planck equation. In the stationary, horizontally homogeneous case adopted here (so that P_E depends only on height z), they can be written as

$$a_i = \frac{\phi_i}{P_E} + \frac{\frac{1}{2}C_0\varepsilon\frac{\partial P_E}{\partial u_i}}{P_E}, \tag{A.9}$$

$$b_i = \sqrt{C_0\varepsilon}, \tag{A.10}$$

where ε denotes the turbulent kinetic energy dissipation rate and ϕ_i is also computed from P_E and are described further in [79, 86, 88, 100].

The two LS model formulations differ in their assumptions for P_E and the resulting expressions for the Langevin coefficients. These specifics are covered in the following sections.

A.2 Surface Layer (SL) Formulation

The SL model incorporates a Gaussian pdf with jointly Gaussian velocity components (u, v, and w), as described in [31, 79, 88], giving a P_E,

$$P_E = \frac{1}{(2\pi)^{3/2}\det\tau_{ij}^{1/2}}\exp\left(-\frac{1}{2}(u_i - U_i)\tau_{ij}^{-1}(u_j - U_j)\right). \tag{A.11}$$

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) [31].

$$a_u = \frac{b_u^2}{A}\left(\overline{u'w'}\,w' - \sigma_w^2 u'\right) + \frac{1}{2}\frac{\partial \overline{u'w'}}{\partial z} + \frac{1}{A}\left(\sigma_w^2 \frac{\partial \sigma_u^2}{\partial z} u'w' - \overline{u'w'}\frac{\partial \sigma_u^2}{\partial z} w'^2 - \overline{u'w'}\frac{\partial \overline{u'w'}}{\partial z} u'w' + \sigma_u^2 \frac{\partial \overline{u'w'}}{\partial z} w'^2\right), \tag{A.12}$$

$$a_v = -\frac{1}{2} b_v^2 \frac{v'}{\sigma_v^2} + \frac{1}{2}\frac{\partial \sigma_v^2}{\partial z}\frac{v'w'}{\sigma_v^2}, \tag{A.13}$$

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

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

A.3 Convective Boundary Layer (CBL) Formulation

The CBL model was developed by Luhar et al. (1989) [86], 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 (A.16),

$$\begin{aligned} P_E &= AP_A + BP_B, \\ P_A &= \frac{1}{\sqrt{2\pi}\sigma_A}\exp\left(\frac{-(w' - \overline{w}_A)^2}{2\sigma_A^2}\right), \\ P_B &= \frac{1}{\sqrt{2\pi}\sigma_B}\exp\left(\frac{-(w' + \overline{w}_B)^2}{2\sigma_B^2}\right). \end{aligned} \tag{A.16}$$

It is extended to 3 dimensions based on Luhar (2002) [54, 119], which takes the horizontal

and vertical velocity fluctuations to be independent. The Langevin coefficients then become,

$$a_w = \frac{\phi}{P_E} - \frac{\frac{1}{2}C_0\varepsilon Q}{P_E}, \tag{A.17}$$

$$a_v = \frac{-v'C_0\varepsilon}{2\sigma_v^2}, \tag{A.18}$$

$$a_u = \frac{-u'C_0\varepsilon}{2\sigma_u^2}. \tag{A.19}$$

The ϕ term has been adapted to heavy particles in Boehm et al. (2005) [100]. 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) [87]. These are functions of the wind velocity profiles, which vary with height and are described in Appendix B. Although the original CBL LS model [86, 87] was a one-dimensional model intended for the well-mixed boundary-layer, Boehm et al. (2008) [54] incorporated wind statistics into this model which transition smoothly from the surface layer to the convective boundary layer above.

Appendix B

Wind statistics profiles by chapter

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.

B.1 Chapter 1: Stable and Unstable Wind Statistics Applied to the Full Boundary Layer

In chapter 1, we applied the 2D SL and CBL formulations to stable and unstable conditions respectively. The wind statistics profiles used for each are included below with their sources for reproducibility.

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 [108] with the stability correction function, ψ_M.

$$\overline{U} = \frac{u_*}{0.4}\left[\ln\left(\frac{z}{z_0}\right) + \psi_M\right]. \tag{B.1}$$

For stable conditions, we use the stability function as reported in Beljaars & Holtslag (1991)[80], 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) [174] also compared various stable wind profiles, including the one presented below, and show that it performs well up to 200 meters above the surface.

$$\psi_M = a\frac{z}{L} + b\left(\frac{z}{L} - \frac{c}{d}\right)\exp\left(-d\frac{z}{L}\right) + \frac{bc}{d}. \tag{B.2}$$

For unstable conditions, we use the stability function given by Paulson (1970) [162]. This has previously been used for other unstable LS simulations [31, 54], and is considered to approximate measurements well [80].

$$\psi_M = -2\ln\left(\frac{1+\alpha}{2}\right) - \ln\left(\frac{1+\alpha^2}{2}\right) + 2\tan^{-1}(\alpha) - \frac{\pi}{2}, \tag{B.3}$$

where,

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

Horizontal wind velocity variance

In stable conditions, we use the following relationship from Kantha and Clayson for the horizontal velocity variance (2000)[82],

$$\sigma_u^2 = 4{u_*}^2\left(1 - \frac{z}{z_i}\right)^{3/2}. \tag{B.5}$$

In unstable conditions, we use the following parameterization from Luhar et al. (2002) [119] for the horizontal wind velocity variance.

$$\sigma_u^2 = (0.6w_*)^2. \tag{B.6}$$

Vertical wind velocity variance

In stable conditions, for the vertical wind velocity variance, we use a relationship from Kantha and Clayson [82],

$$\sigma_w^2 = 3{u_*}^2\left(1 - \frac{z}{z_i}\right)^{3/2}. \tag{B.7}$$

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)[54]. 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.

$$\sigma^2_{w,CBL} = 1.7{w_*}^2(z/z_i)^{2/3}(1-0.9z/z_i)^{4/3}, \tag{B.8a}$$

$$\sigma^2_{w,neutral} = {u_*}^2(1.7 - z/z_i), \tag{B.8b}$$

$$\sigma^2_{w,merged} = \frac{(1-\exp(z/L)){w_*}^3\sigma^2_{w,CBL} + 5\exp(z/L){u_*}^3\sigma^2_{w,neutral}}{(1-\exp(z/L)){w_*}^3 + 5\exp(z/L){u_*}^3}. \tag{B.8c}$$

$$\tag{B.8d}$$

Velocity covariance

In unstable conditions, the covariance between downwind and vertical velocities is required. We use the following parameterizations used in Aylor and Flesch (2001) [31],

$$\overline{u'w'} = -u_*^2 \tag{B.9}$$

Lagrangian Timescale

In all stabilities, we compute the Lagrangian time scale using[31, 87],

$$\tau = \frac{2\sigma_w^2}{C_0\varepsilon}. \tag{B.10}$$

Turbulence kinetic energy dissipation rate

In stable conditions, we use the profile suggested by Rodean (1996) [79] for the entire stable boundary layer,

$$\varepsilon = \frac{{u_*}^3}{0.4 * z}\left(1 + 3.5\frac{z}{L}\right)\left(1 - 0.85 * \frac{z}{z_i}\right)^{3/2}. \tag{B.11}$$

Rodean (1996) discusses that this profile was formed by fitting to a second-order turbulence model[175], 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) [54] to LS modeling, and found previously using Large Eddy Simulations[157],

$$\varepsilon = 0.4\frac{{w_*}^3}{z_i} + \frac{{u_*}^3(1 - z/z_i)}{0.4z(1 - 15 * z/L)^{1/4}}. \tag{B.12}$$

B.2 Chapter 2: Unstable Wind Statistics Applied to the Surface layer

In chapter 2, we applied the 3D SL formulation to unstable daytime conditions to simulate near-source dispersal. The wind statistics profiles used are included below for reproducibility, and come entirely from Aylor and Flesch (2001) [31].

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 [108] with the stability correction function, ψ_M.

$$\overline{U} = \frac{u_*}{0.4}\left[\ln\left(\frac{z}{z_0}\right) + \psi_M\right]. \tag{B.13}$$

We use the stability function given by Paulson (1970) [162]. This has previously been used

for other unstable LS simulations [31, 54], and is considered to approximate measurements well [80].

$$\psi_M = -2\ln\left(\frac{1+\alpha}{2}\right) - \ln\left(\frac{1+\alpha^2}{2}\right) + 2\tan^{-1}(\alpha) - \frac{\pi}{2}, \tag{B.14}$$

where,

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

Horizontal wind velocity variance

We use the following parameterizations used in Aylor and Flesch (2001) [31] for the horizontal wind velocity variances,

$$\sigma_u^2 = \sigma_v^2 = u_*^2(4 + 0.6 * (-z_i/L)^{2/3}. \tag{B.16}$$

Vertical wind velocity variance

In unstable conditions, we use the following parameterization used in Aylor and Flesch (2001) [31],

$$\sigma_w^2 = 1.56u_*^2(1 - 3 * (z/L)^{2/3}. \tag{B.17}$$

Vertical wind velocity variance

We use the following parameterization used in Aylor and Flesch (2001) [31],

$$\sigma_w^2 = 1.56u_*^2(1 - 3 * (z/L)^{2/3}. \tag{B.18}$$

Velocity covariance

For this model, the covariance between downwind and vertical velocities is included. We use the following parameterizations used in Aylor and Flesch (2001) [31],

$$\overline{u'w'} = -u_*^2 \tag{B.19}$$

Lagrangian Timescale

We compute the Lagrangian time scale as follows with the timescale correction as written in Aylor and Flesch (2001) [31], where they take β to be 1.5.

$$T_L = \frac{0.5z}{\sigma_w}\left(1 - 6\frac{z}{L}\right)^{1/4}, \tag{B.20}$$

$$\tau = fT_L \qquad f = \frac{1}{\sqrt{1 + (\beta v_s/\sigma_w)^2}}, \tag{B.21}$$

Turbulence kinetic energy dissipation rate

We compute the turbulence kinetic energy dissipation rate using[31, 87],

$$\varepsilon = \frac{2\sigma_w^2}{C_0\tau}. \tag{B.22}$$

B.3 Chapter 3: Unstable wind statistics Applied to the Full Boundary Layer

In chapter 3, we applied the 3D SL and CBL formulations to unstable daytime conditions to simulate disperal within about 1 km of the source. The wind statistics profiles used for each are included below with their sources for reproducibility.

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 [108] with the stability correction function, ψ_M.

$$\overline{U} = \frac{u_*}{0.4}\left[\ln\left(\frac{z}{z_0}\right) + \psi_M\right]. \tag{B.23}$$

For unstable conditions, we use the stability function given by Paulson (1970) [162]. This has previously been used for other unstable LS simulations [31, 54], and is considered to approximate measurements well [80].

$$\psi_M = -2\ln\left(\frac{1+\alpha}{2}\right) - \ln\left(\frac{1+\alpha^2}{2}\right) + 2\tan^{-1}(\alpha) - \frac{\pi}{2}, \tag{B.24}$$

where,

$$\alpha = (1 - 15\frac{z-d}{L})^{1/4}. \tag{B.25}$$

Horizontal wind velocity variance

We use the following parameterization from Luhar et al. (2002) [119] for the horizontal wind velocity variance.

$$\sigma_u^2 = (0.6 w_*)^2. \tag{B.26}$$

Vertical wind velocity variance

We apply the merged parameterization from Boehm et al. (2005)[54]. 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.

$$\sigma_{w,CBL}^2 = 1.7 {w_*}^2 (z/z_i)^{2/3} (1 - 0.9 z/z_i)^{4/3}, \tag{B.27a}$$

$$\sigma_{w,neutral}^2 = {u_*}^2 (1.7 - z/z_i), \tag{B.27b}$$

$$\sigma_{w,merged}^2 = \frac{(1 - \exp(z/L)) {w_*}^3 \sigma_{w,CBL}^2 + 5 \exp(z/L) {u_*}^3 \sigma_{w,neutral}^2}{(1 - \exp(z/L)) {w_*}^3 + 5 \exp(z/L) {u_*}^3}. \tag{B.27c}$$

$$\tag{B.27d}$$

Velocity covariance

In unstable conditions, the covariance between downwind and vertical velocities is required. We use the following parameterization described in Rodean (1996) [79, 81],

$$\overline{u'w'} = -u_*^2 * (1 - z/z_i)^{3/2-q}, \tag{B.28}$$

$$q = \frac{-0.5 * z_i/L}{1 - z_i/L}. \tag{B.29}$$

$$\tag{B.30}$$

Lagrangian Timescale

We compute the Lagrangian time scale using[31, 87],

$$\tau = \frac{2\sigma_w^2}{C_0\varepsilon}. \tag{B.31}$$

Turbulence kinetic energy dissipation rate

In stable conditions, we use the profile suggested by Rodean (1996) [79] for the entire stable boundary layer,

$$\varepsilon = \frac{{u_*}^3}{0.4 * z}\left(1 + 3.5\frac{z}{L}\right)\left(1 - 0.85 * \frac{z}{z_i}\right)^{3/2}. \tag{B.32}$$

Rodean (1996) discusses that this profile was formed by fitting to a second-order turbulence model[175], 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) [54] to LS modeling, and found previously using Large Eddy Simulations[157],

$$\varepsilon = 0.4\frac{{w_*}^3}{z_i} + \frac{{u_*}^3(1 - z/z_i)}{0.4z(1 - 15 * z/L)^{1/4}}. \tag{B.33}$$

Appendix C

Supplementary Figures For Chapter 1

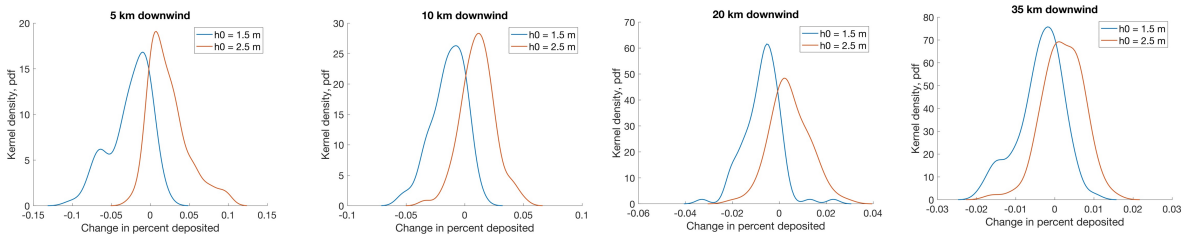
Supplementary Figure S1

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 C.1(A) shows pdfs of the change in percent-deposited at 5, 10, 20, and 35 km downwind, while Figure C.1(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 C.1(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)

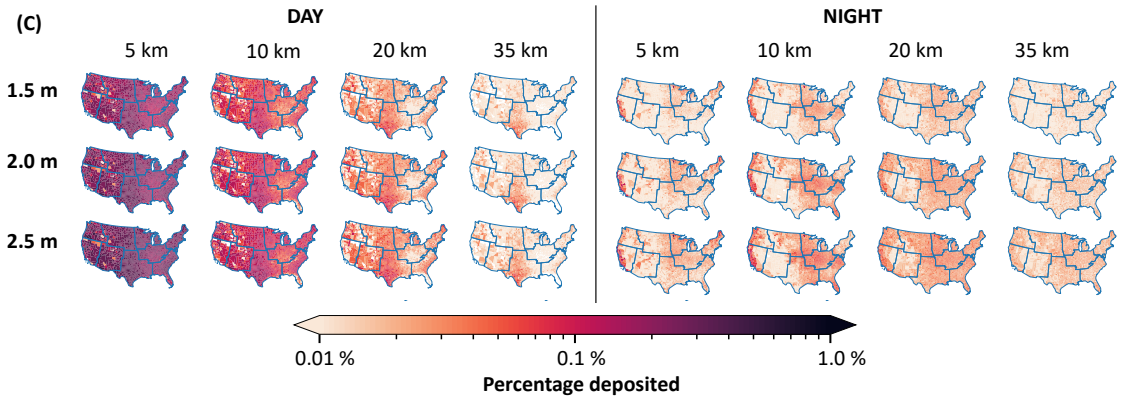


Figure C.1: Sensitivity analysis conducted by varying the release height, where 2 m is the height originally used in the paper. (A) pdfs of the change (from a 2-m release height) in percent-deposited at 5, 10, 20, and 35 km downwind for both day and night conditions together. (B) Mean and median change 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

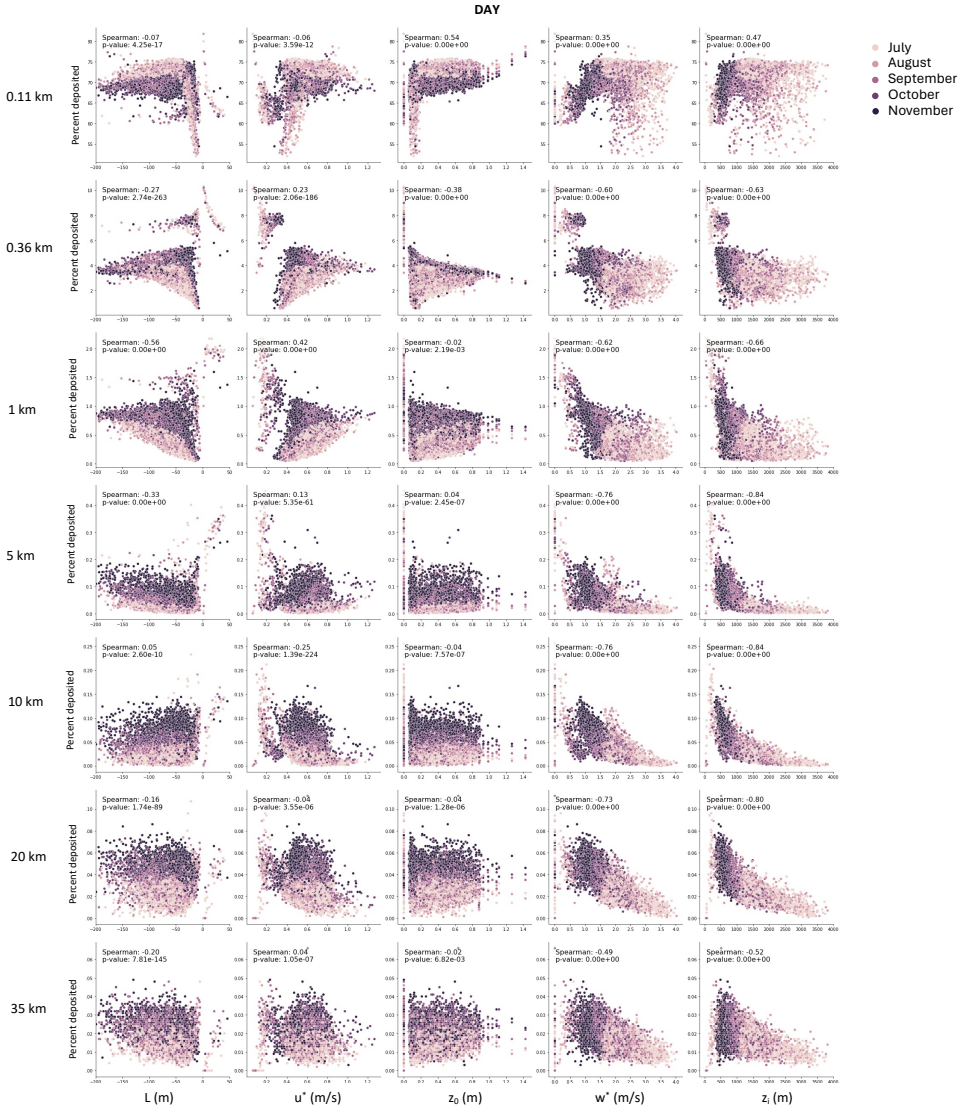


Figure C.2: Scatterplots of five meteorological parameters for all day simulations vs. the percentage of particles deposited at distances downwind of the source. The Spearman correlation coefficients relating depositions at each downwind distance with the respective meteorological parameter are denoted 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

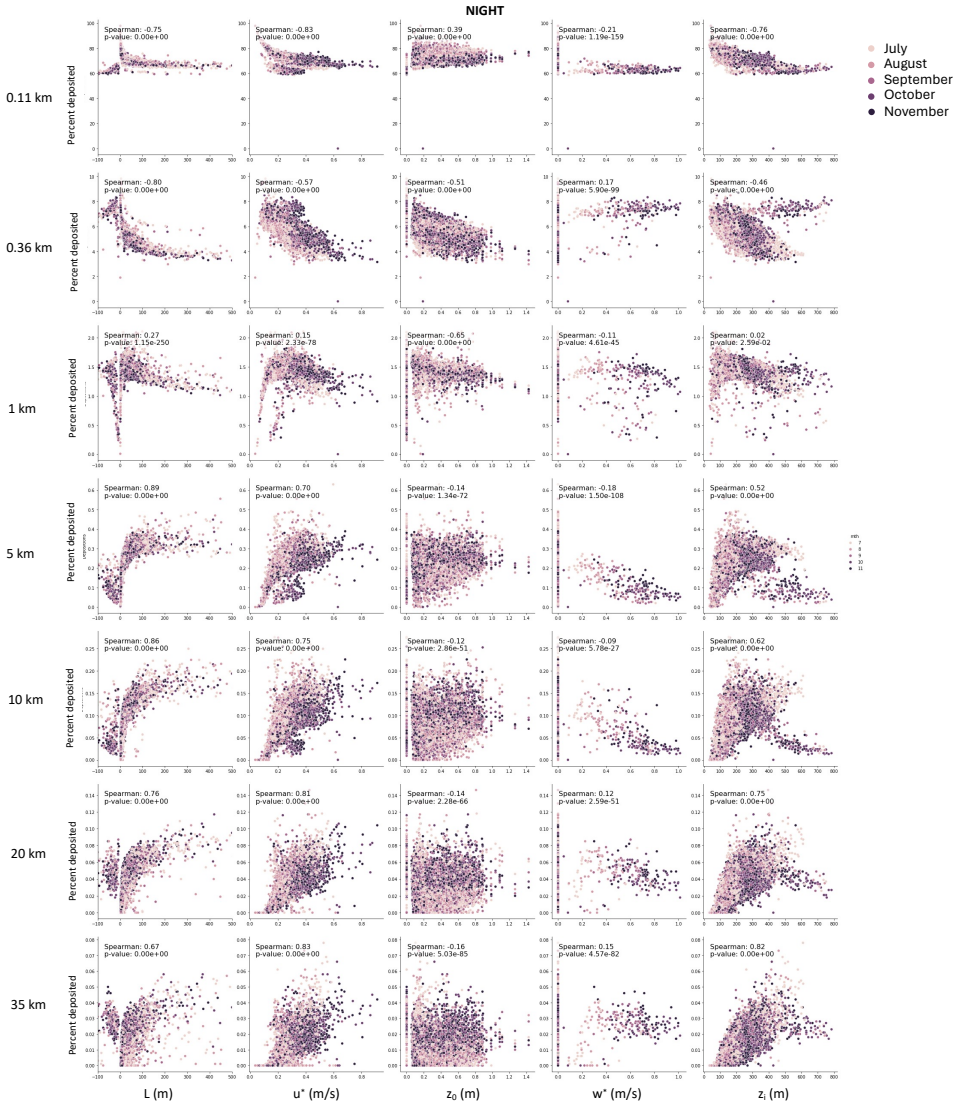


Figure C.3: Scatter plots of meteorological input parameters vs. depositions for night cases. Scatterplots of five meteorological parameters for all night simulations vs. the percentage of particles deposited at distances downwind of the source. The Spearman correlation coefficients relating depositions 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

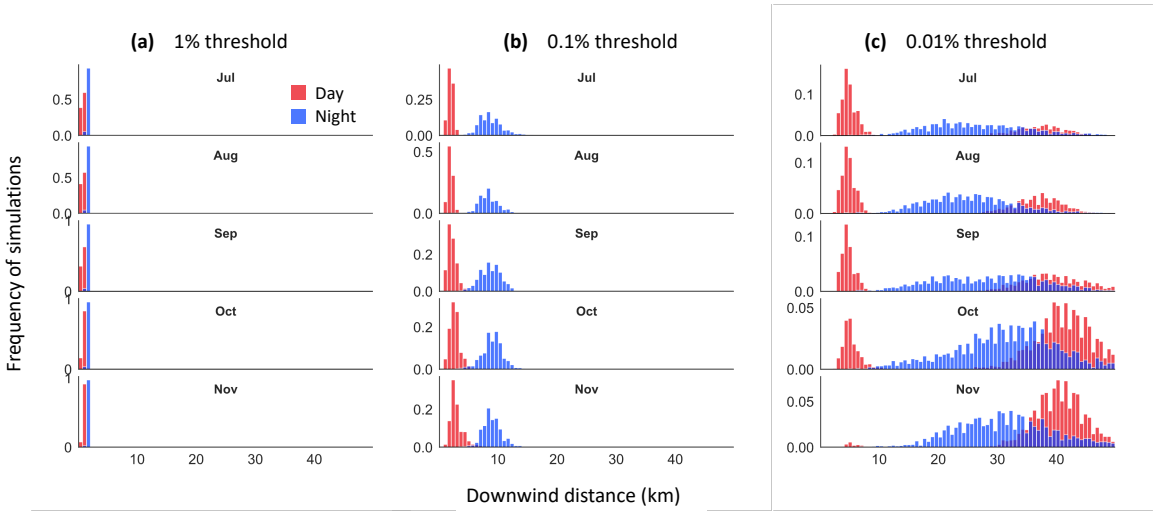


Figure C.4: Distances at which dispersal kernels first fall below a threshold for each month: (a) 1%, (b) 0.1%, and (c) 0.01%. Red represents day simulations, while blue represents night. Seasonal variation is most pronounced for the 0.01% threshold distances, where the frequency of daytime distances beyond 30 km progressively increases from July to November.

Supplementary Figure S5

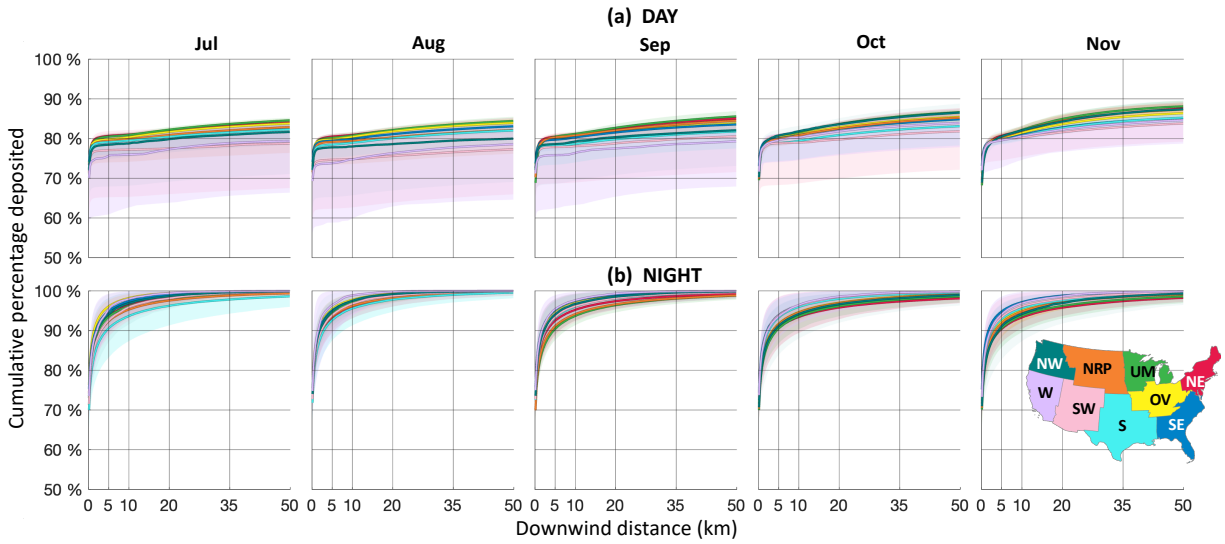


Figure C.5: Median cumulative depositions for each month during (a) day and (b) night, separated by US climate region: Northeast (NE), Upper Midwest (UM), Ohio Valley (OV), Southeast (SE), Northern Rockies & Plains (NRP), South (S), Southwest (SW), Northwest (NW), and West (W). Shading represents data 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

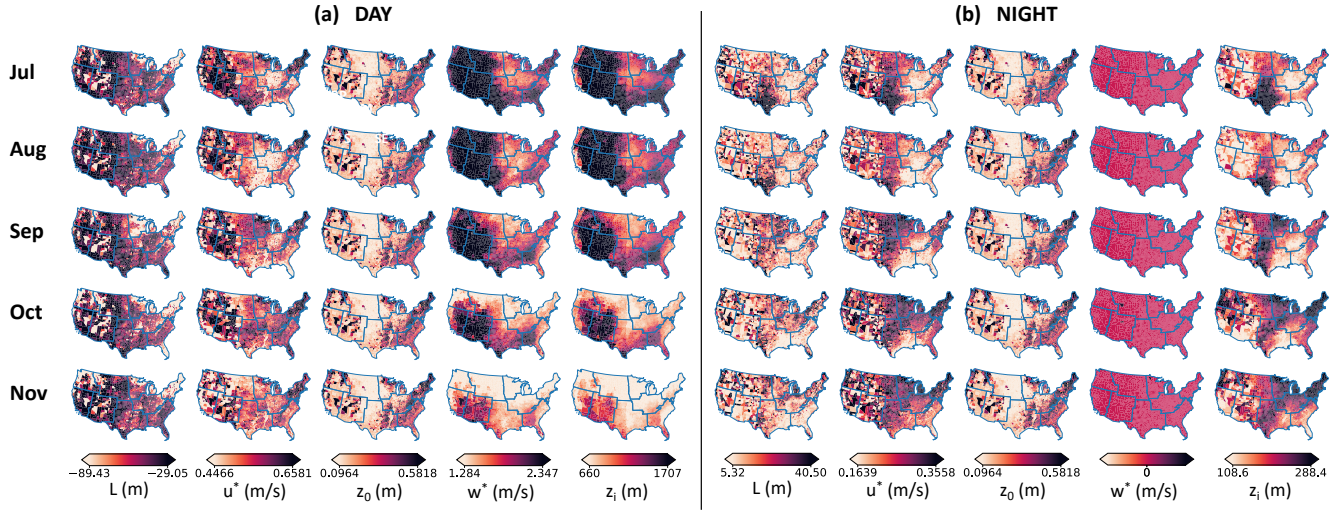


Figure C.6: Heat maps of five meteorological parameters for all (a) daytime and (b) nighttime simulations over five months from July to November. The spatial and seasonal patterns visualized here mirror the deposition patterns shown in the main paper. During the day, the Southwest region maintains the 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.

C.1 Supplementary Figure S7

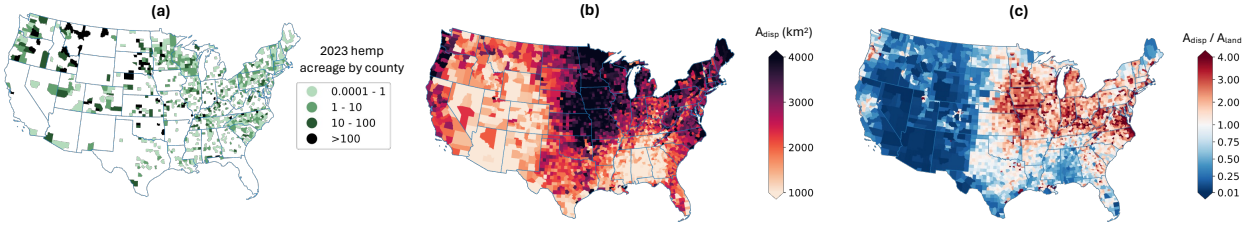


Figure C.7: Components of the vulnerability metric. (a) The acreage of hemp, A_{hemp}, planted in each county as of 2023 [2], where darker colors indicate greater planted hemp acreage. (b) The dispersal area A_{disp}, or area within a circle of radius equal to the average 0.01%-threshold distance, where 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 $A_{disp} > A_{land}$.