Article

Not peer-reviewed version

Mapping High-Risk Disease Zones in Strawberry Fields Using Drone Imagery and Random Survival Forests

Jon R. Detka *, Adam J. Purdy, Forrest S. Melton, Oleg Daugovish, Christopher A. Greer, Frank N. Martin

Posted Date: 12 February 2026

doi: 10.20944/preprints202602.1030.v1

Keywords: strawberry production; canopy health; machine learning; multispectral imaging; random survival forest; risk analysis; uncrewed aerial system (UAS)

Article

Mapping High-Risk Disease Zones in Strawberry Fields Using Drone Imagery and Random Survival Forests

Jon R. Detka [1,*], Adam J. Purdy [1,2], Forrest S. Melton [2], Oleg Daugovish [3], Christopher A. Greer [4] and Frank N. Martin [5]

1 California State University, Monterey Bay. Seaside, CA, USA
2 NASA Ames Research Center, Moffett Field, CA, USA
3 University of California, Cooperative Extension, Ventura, CA, USA
4 University of California, Cooperative Extension, San Luis Obispo, CA, USA
5 U.S. Department of Agriculture, Agricultural Research Services, Salinas, CA, USA
* Correspondence: jdetka@csumb.edu

Abstract

Early detection of canopy decline in strawberry production is essential for timely management, yet visual scouting often misses subtle or spatially heterogeneous symptoms. We developed a UAV-based monitoring framework that integrates multispectral imagery, plant-level canopy metrics, clustering, and Random Survival Forest (RSF) modeling. This framework was used to predict the onset and spatial progression of soilborne pathogen-associated canopy decline in three commercial strawberry fields in Oxnard, California. Nine UAV surveys collected from December 2022 to June 2023 were processed into 159,220 plant-level monitoring units. NDRE- and Redness Index–based classifications quantified proportional and absolute canopy dieback within standardized hexagonal units and supported a time-to-event modeling approach. RSF models achieved consistently high concordance during periods of active decline, with strongest performance in the field exhibiting the greatest disease pressure. Spatial risk maps revealed early hotspots that expanded into contiguous high-risk zones by June, while fields with minimal visible symptoms showed diffuse but consistent risk patterns. Post-hoc comparison with operational fumigation rates (280, 336, and 392 kg Pic-Clor 60/ha) showed no consistent association with predicted canopy risk, consistent with the possibility that lower application rates may be sufficient in portions of fields with historically low disease pressure. These results demonstrate that UAV multispectral time series combined with survival modeling can track fine-scale spatiotemporal canopy decline and provide an early-warning framework to support spatially targeted disease monitoring and management in commercial strawberry systems.

Keywords: strawberry production; canopy health; machine learning; multispectral imaging; random survival forest; risk analysis; uncrewed aerial system (UAS)

---

1. Introduction

Strawberries are one of the most widely cultivated fruit crops worldwide, grown commercially in 76 countries across Asia, North and Central America, and North Africa, with global demand continuing to rise [1]. In 2023, the United States ranked second in global strawberry production, yielding approximately 1.25 million metric tons, behind only China at 4.2 million metric tons [2]. Within the U.S., California produces nearly 90% of the nation's strawberries and was ranked as the fourth largest strawberry producer globally [3,4]. The crop is a vital contributor to both national fruit supply and the state's agricultural economy, with an estimated farmgate value of $2.97 billion and 40,285 acres in production as of 2023 [3,5].

Crown and root rot diseases remain among the most persistent and economically damaging challenges in California strawberry production. The primary soilborne pathogens include *Macrophomina*

doi:10.20944/preprints202602.1030.v1

phaseolina, *Fusarium oxysporum* f. sp. *fragariae*, *Verticillium dahliae*, and *Phytophthora cactorum* [6]. All of these pathogens persist in soil and can accumulate high inoculum levels over time, complicating long-term disease management in commercial fields [7,8]. Symptoms associated with these pathogens vary by disease type but commonly include crown necrosis, vascular discoloration, stunting, chlorosis, and eventual canopy collapse. In all cases, visible symptoms typically emerge only after substantial root or vascular damage has occurred, making early detection difficult [9]. These symptoms are also easily confused with those of other soilborne pathogens, which complicates diagnosis and delays intervention. As a result, high-risk zones are often not identified until plant decline is well advanced, increasing the likelihood of transplant failure and yield loss [9]. In the context of soilborne pathogens, early detection remains valuable for identifying emerging high-risk zones and informing spatially targeted management actions (e.g., irrigation and stress mitigation) as well as planning decisions for subsequent production seasons.

Although strawberry is a perennial species, commercial production in California is managed on an annual planting cycle, with fields replanted each year, often repeatedly in the same soil. As a result, soilborne pathogens such as *M. phaseolina* pose ongoing threats to future plantings due to their long-term survival in soil and the potential for inoculum persistence across seasons. Although *M. phaseolina* is generally considered a broad-host-range pathogen, the strawberry-aggressive genotype present in California may exhibit greater host specificity, and its role in disease carryover across crop rotations remains an active area of investigation. Recent surveys in California indicate that the prevalence and severity of crown and root rot diseases vary across regions and seasons, with several pathogens exhibiting increased severity under warmer production conditions. In coastal production regions such as Oxnard, Fusarium wilt has become an increasing concern, including reports of novel races capable of overcoming single-gene resistance in current cultivars [10–12]. In this context, early detection remains valuable for identifying emerging high-risk zones and informing spatially targeted management actions (e.g., irrigation and stress mitigation) as well as planning decisions for subsequent production seasons. Together, these dynamics underscore the need for spatially explicit disease risk mapping to inform planting, fumigation, and in-field management decisions.

To manage these threats, growers typically rely on pre-plant soil fumigation with broad-spectrum fumigants such as chloropicrin–1,3-dichloropropene formulations, often applied uniformly across entire fields. While these broad-spectrum fumigants are effective, fumigation imposes additional costs, requires specialized application, and is increasingly subject to regulatory scrutiny [13]. Additionally, fumigation efficacy can be limited in fields with uneven pathogen distribution or high baseline inoculum levels, often resulting in patchy stand establishment and midseason plant decline despite treatment.

Current field monitoring practices offer limited ability to detect disease onset early or track its progression over time. Ground scouting and end-of-season yield assessments typically capture symptoms only after significant physiological damage and yield loss have occurred. While these methods can document disease impact retrospectively, they rarely provide the spatial extent or temporal resolution needed to guide timely, in-season management decisions. Field studies similarly emphasize late-stage disease ratings or final harvest outcomes, which are useful for evaluating treatment efficacy but poorly suited to identifying when and where symptoms first emerge. As a result, opportunities for early risk identification, targeted scouting, or strategic management planning remain limited, particularly for pathogens with few effective in-season control options [14].

There is a growing interest in using high-resolution remote sensing tools together with precision disease management strategies to target high-risk zones. Such tools would improve scouting efficiency, and support more refined risk stratification within fields, particularly in production systems where disease pressure is known or expected.

Recent advances in remote sensing and statistical modeling offer promising tools for disease monitoring and modeling plant health at the scales necessary to improve management and disease mitigation [15,16]. Uncrewed aerial vehicles (UAVs), or drones, enable frequent, field-scale collection of

 doi:10.20944/preprints202602.1030.v1

high-resolution imagery, allowing for repeatable, non-destructive monitoring of plant canopy condition throughout the growing season. When equipped with multispectral or hyperspectral sensors, imagery generated from UAVs can detect subtle changes in canopy vigor that may signal physiological stress well before visible symptoms appear [17–19]. Onboard sensors capture spectral patterns that, when paired with ground validation, can be used to infer zones of stress potentially associated with disease or other agronomic factors [20]. While remote sensing tools are increasingly promoted for early disease detection, it is important to recognize that imagery collected from UAVs detects canopy-level stress responses, not pathogens themselves. Over-interpreting spectral signals without adequate field validation has been noted as a risk in the literature [21].

Vegetation indices such as the Normalized Difference Vegetation Index (Eq. 1), Normalized Difference Red Edge Index (Eq. 2), and Redness Index (Eq. 3) have been shown to correlate with early signs of chlorosis, canopy thinning, and dieback in a range of crops, including strawberry [21–23]. Research has demonstrated that NDRE and NDVI are sensitive to foliar symptoms of *Mycosphaerella fragariae* in strawberry [24]. However, most prior optical and spectral approaches in strawberry have focused on foliar diseases at the leaf level rather than the canopy scale, where visible symptoms such as discoloration or necrosis, and pathogen signs, drive spectral change [25–27]. In contrast, our study explores the use of these indices to detect canopy stress associated with soilborne pathogens such as *M. phaseolina*. In this context, spectral signals are indirect and reflect secondary physiological responses like stunting, reduced canopy vigor, and dieback, rather than direct pathogen presence. This distinction underscores the need to pair canopy-level imagery with spatiotemporal modeling frameworks that can track disease progression under real-world field conditions.

Modeling disease progression based on canopy metrics in real-world field systems presents several challenges. Stress responses are often nonlinear, influenced by spatial and temporal context, and not all plants exhibit visible decline during the observation window, resulting in censored outcomes. Plants that remain healthy through the final survey contribute censored outcomes, meaning the exact time of potential decline is unknown. Traditional regression approaches struggle to accommodate these complexities. At the same time, advances in machine learning have introduced new tools for modeling complex biological processes, including disease development in crops [28,29]. Among these, Random Survival Forest (RSF) provides a powerful, nonparametric framework for time-to-event prediction. RSFs can model nonlinear relationships, handle censored data, and integrate spatiotemporal predictors—making them well-suited for analyzing plant-level trajectories of canopy decline under field conditions [30,31].

Here, we evaluate whether drone-based multispectral imagery, when paired with spatial and temporal modeling of plant-level canopy metrics, can be used to detect and predict early-stage canopy decline in strawberry fields with persistent *M. phaseolina* pressure. Using data from three commercial fields in Oxnard, California, we assess the predictive performance of Random Survival Forest models trained on field-derived vegetation indices and spatial context metrics. We interpret the results in the context of spatially targeted disease management and conclude by discussing the implications and practical applications of this approach across spatiotemporal scales.

2. Methods

2.1. Study Site

The study was conducted at a commercial strawberry farm in Oxnard, California. UAV flight surveys focused on 39-40 bed rows across three adjacent fields, each approximately 0.8 ha in size (Figure 1). The fields were planted with the 'Fronteras' strawberry variety, installed in late November 2022, and arranged in bed rows four plants wide. 'Fronteras' is a short-day cultivar developed by the University of California, Davis strawberry breeding program. It is widely grown in California's coastal production regions, particularly Ventura, Santa Barbara, and Monterey counties due to its early fruiting, high yield potential, and firm fruit that is well-suited for commercial packing. The cultivar

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 12 February 2026 doi:10.20944/preprints202602.1030.v1

shows resistance to Fusarium wilt, moderate resistance to Verticillium wilt and Phytophthora root rot, and moderate susceptibility to *M. phaseolina* (California Strawberry Commission, 2024).

Prior to planting, the site was treated with Pic-Clor 60, a pre-plant soil fumigant containing chloropicrin and 1,3-dichloropropene that is widely used in California strawberry production for broad-spectrum suppression of soilborne pathogens. The fumigant was applied before bed formation and sealed with a totally impermeable film (TIF) to enhance fumigant retention and efficacy. The three fields also contained grower-applied pre-plant fumigation blocks with variable rates based on historic disease pressure (Figure 1); these operational treatment zones were incorporated later in a post-hoc analysis to evaluate whether RSF-derived canopy risk differed systematically among fumigation rates.

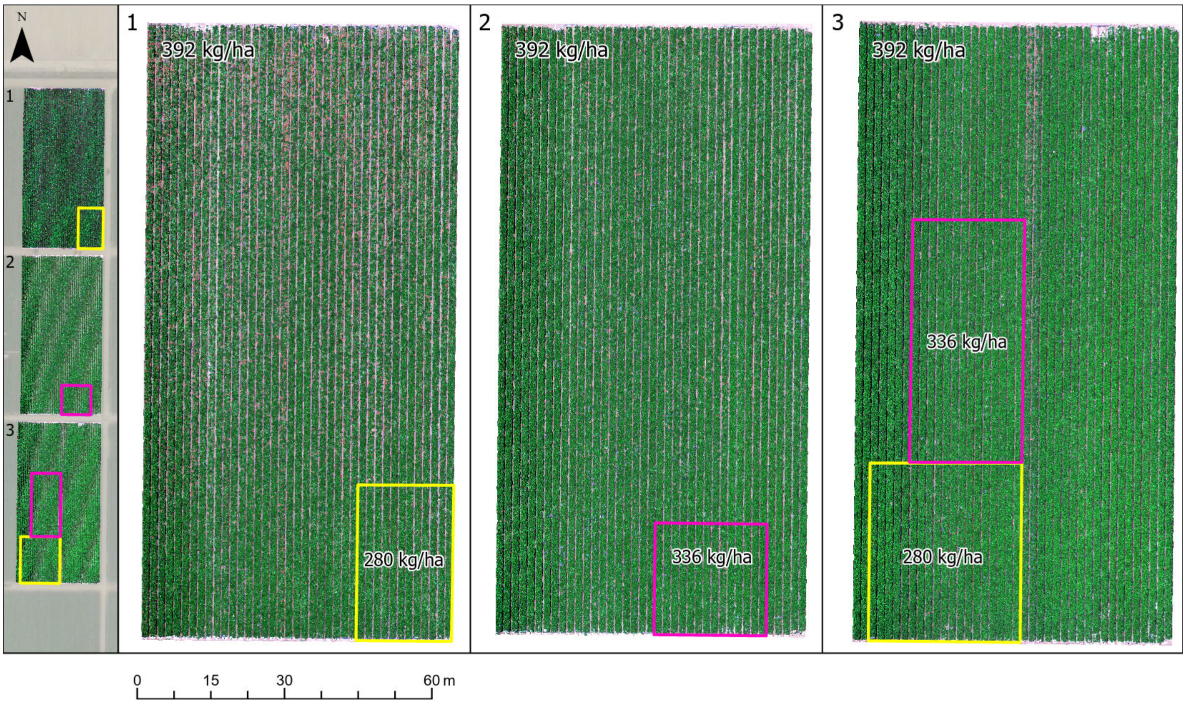

Figure 1. Spatial arrangement of fumigation treatments across the three study fields. Background shows the June 20, 2023 UAV RGB orthomosaic. Colored outlines indicate variable Pic-Clor 60 soil fumigation rates: yellow = Low 280 kg/ha (250 lbs/acre), magenta = Medium 336 kg/ha (300 lbs/acre). All other areas received the standard rate of 392 kg/ha (350 lbs/acre).

2.2. UAV Data Collection and Imagery Processing

2.2.1. UAV Data Collection

Aerial surveys were conducted using a Matrice 200 V2 quadcopter (Da-Jiang Innovations Sciences and Technologies Ltd. (DJI), Shenzhen, China). Multispectral image data were collected using the MicaSense Altum-PT sensor (EagleNXT, Allen, TX, USA), which measures radiance from five spectral bands (red, green, blue, near-infrared, and red-edge) and a longwave infrared thermal band. To support radiometric correction, the Altum-PT was paired with an irradiance light sensor (Downwelling Light Sensor (DLS 2), EagleNXT, Allen, TX, USA) that recorded sun angle and ambient light conditions during each flight.

To support radiometric calibration, reflectance calibration images were collected using the factory-supplied MicaSense Altum-PT calibration panel (Panel 2, EagleNXT, Allen, TX, USA) prior to each flight. Panel images were used during post-processing to convert raw digital numbers to surface reflectance and, in combination with data from the sensor's downwelling light sensor (DLS 2), to improve radiometric consistency across survey dates.

Aerial surveys were conducted at 30 m above ground level (AGL) with 85% forward and 75% side overlap with flight speeds below 2 m/s to minimize motion blur. All flights were conducted within one hour of local solar noon to minimize shadow effects and variability in solar illumination. A total of nine drone flights were conducted over three adjacent strawberry fields from December 2022 to June 2023. Flights were spaced approximately nine weeks apart in winter (December 7, February 9), and 1–3 weeks apart in spring (March 28, April 21, May 15, May 22) and early summer (June 1, June 10, June 20).

To ensure accurate georeferencing, seven ground control points (GCPs) were deployed throughout each field. Each GCP consisted of a 0.25 m × 0.25 m black-and-white checkerboard target. The center of each GCP was surveyed using an REACH RS+ single-band real-time kinematic (RTK) GNSS rover (Emlid Tech Kft., Hungary) connected via an NTRIP correction service, providing horizontal accuracy of approximately ±0.03 m and vertical accuracy of ±0.05 m.

2.2.2. Imagery Processing

Imagery was processed using Pix4Dmapper structure-from-motion photogrammetry software (v4.9.0; Pix4D SA, Lausanne, Switzerland). Each survey yielded approximately 2,000 individual images across 5 bands per 1 hectare. Processing followed Pix4D's standard workflow, which includes image alignment, sparse and dense point cloud generation, digital surface model (DSM) creation, and generation of georeferenced orthomosaics and reflectance maps. Ground control points (GCPs) were manually marked across overlapping images to improve georeferencing accuracy.

Image alignment was optimized using Pix4D's agricultural processing settings to improve feature matching in vegetated fields with repetitive textures. Radiometric correction was applied using both the Altum-PT's downwelling light sensor (DLS 2) and reference images of the custom-built reflectance calibration target, which was imaged prior to each flight. These corrections ensured consistent surface reflectance values across survey dates and varying illumination conditions within each flight.

2.3. Plant-Level Canopy Metric Extraction

2.3.1. Feature Extraction and Classification

Post-processed reflectance rasters were analyzed in ArcGIS Pro (v3.3.1; Esri Inc., Redlands, CA, USA) using Python scripts built on the ArcPy library. For each survey date, multispectral composites were clipped by field and vegetation indices were computed from individual band rasters. The indices included the Normalized Difference Vegetation Index (NDVI; [32]) for canopy vigor:

$$\mathrm{NDVI} = \frac{NIR - Red}{NIR + Red} \tag{1}$$

the Normalized Difference Red Edge Index (NDRE; [33]) for early detection of stress in maturing plants:

$$\mathrm{NDRE} = \frac{NIR - RedEdge}{NIR + RedEdge} \tag{2}$$

and a normalized red–green difference index (RI), mathematically equivalent to the normalized green–red difference index (NGRDI; [34]) but with reversed sign, which captures shifts in red and green reflectance associated with senescence or canopy decline:

$$\mathrm{RI} = \frac{R - G}{R + G} \tag{3}$$

NDVI was used to generate binary canopy masks that were converted to polygons and assigned to bed rows via spatial buffering. Polygon centroids were extracted from these canopy objects, deduplicated by proximity, and assigned unique plant-level identifiers for downstream analysis.

Hexagonal buffers (0.17 m radius) were generated around each canopy centroid to standardize the spatial unit of analysis. Within each hexagon, NDRE- and RI-based thresholds (Table 1) were applied using a hierarchical, rule-based classification approach to assign pixels as healthy, dieback, or non-vegetated. NDRE was used to identify vegetated canopy, and RI was subsequently used to refine canopy condition within those areas; pixels not meeting any NDRE or RI threshold were assigned to an unclassified state (0) and excluded from analysis to reduce ambiguity associated with mixed pixels and transitional canopy states. Final canopy class assignments were based on the combined response of both indices rather than independent thresholding.

These classifications were summarized into plant-level metrics capturing two complementary aspects of canopy decline. Relative dieback was calculated as the proportion of dieback pixels normalized by the total number of vegetated canopy pixels (healthy + dieback) within each hexagon, providing a within-canopy measure of decline severity independent of plant size. In contrast, absolute dieback was calculated as the proportion of dieback pixels relative to all pixels within the hexagon, including non-vegetated background, thereby capturing the total magnitude of canopy loss associated with stunting or reduced canopy extent. Together, these metrics allowed the model to distinguish plants exhibiting severe proportional decline within an otherwise intact canopy from those experiencing larger overall canopy loss. Plant-level canopy metrics were extracted from hexagon buffers representing individual plants, resulting in 159,220 unique plant records with repeated spectral measurements across survey dates (Table 2).

Threshold values for NDRE and RI were selected empirically based on inspection of index value distributions across flights and visual comparison with canopy condition in the orthomosaics. These thresholds were chosen to consistently separate healthy canopy, declining vegetation, and non-vegetated background while conservatively excluding ambiguous pixels at canopy edges. An intermediate unclassified range was intentionally retained to minimize misclassification of mixed pixels between plant canopy and the surrounding soil or furrow. Classification thresholds were held constant across all fields and survey dates to maximize consistency and minimize mixed-pixel artifacts, which was essential for generating stable plant-level metrics used in survival modeling.

Table 1. Threshold values for Normalized Difference Red Edge Index (NDRE) and Redness Index (RI) used to classify canopy status. NDRE and RI thresholds were applied hierarchically and include intentional gaps. Pixels not meeting any NDRE or RI range were assigned to an unclassified category (0) and excluded from analysis.

Canopy Class	NDRE Range	RI Range	Interpretation
Healthy	≥ 0.60	-1.00 to -0.30	Dense, photosynthetically active canopy
Dieback	0.45–0.59	0.22 to 0.37	Thinning or declining canopy; moderate stress
Non-vegetated	0.14–0.41	-0.01 to 0.16	Bare soil, plastic mulch, or senesced vegetation
Unclassified	Outside ranges	Outside ranges	No clear classification; excluded from analysis

Table 2. Area and number of plant monitoring units within each fumigation rate class by field. Plant counts represent plant-level hexagon units used in RSF analyses. Total values summarize all fumigation rate classes within each field.

Field	Treatment	Fumigation Rate (kg/ha)	Area (ha)	Plant Count
1	Standard	392	0.733	48,318
	Low	280	0.062	4,102
	TOTAL	–	**0.795**	**52,420**
2	Standard	392	0.736	48,830
	Medium	336	0.052	3,455
	TOTAL	–	**0.788**	**52,285**
3	Standard	392	0.597	39,478
	Medium	336	0.113	7,540
	Low	280	0.112	7,497
	TOTAL	–	**0.822**	**54,515**

2.3.2. Critical Cluster Identification and Proximity Analysis

For each flight, plants were assigned a binary "critical" status based on k-means clustering ($k = 3$) applied to standardized plant-level canopy metrics and their temporal changes. The cluster exhibiting the highest dieback severity was designated as the critical group. Local neighborhood effects were quantified as the Euclidean distance to the nearest other critical plant using a *cKDTree* from the *scipy.spatial* module [35]. When multiple critical plants were present, distances were calculated as the nearest-neighbor distance among critical plants. When only a single critical plant was present in a flight, that plant was assigned a distance of zero, and all non-critical plants were assigned the maximum observed distance for that flight to reflect the absence of nearby critical neighbors. To avoid information leakage, the critical cluster was identified per flight using only training plants from an 80/20 train–test split at the individual plant level and then applied to all plants. Both the proximity metric and the per-flight critical indicator were retained as predictors for downstream survival modeling.

2.3.3. Survival Dataset Construction

A survival dataset was created by merging repeated observations for each plant across all flight dates and formatting them into a start–stop interval structure appropriate for time-to-event analysis. The event indicator was defined as whether a plant entered the critical cluster at any flight. Time was computed as the number of days since the plant's first observation, based on the UAV flight calendar. Plants that never entered the critical category during the observation period were treated as right-censored, meaning that the event of interest did not occur before the final survey, and the exact time of potential future decline was unknown.

2.4. Random Survival Forest Modeling

2.4.1. RSF Modeling and Evaluation

Random Survival Forest (RSF) models were implemented in Python using the *scikit-survival* package [36]. Models were fit separately for each field, with performance evaluated using stratified 5-fold cross-validation. Predictor variables included plant-level canopy metrics for proportional dieback, absolute dieback area, and change in canopy condition, and spatial features such as proximity to critical plants and bed-row position. To avoid information leakage, model training and evaluation were conducted using plant-level splits and temporally ordered flights, with critical clusters defined using training plants only.

Model performance was assessed primarily using the concordance index (C-index), which quantifies a model's ability to correctly rank survival risk. The C-index is well-suited for censored time-to-event data because it evaluates whether individuals experiencing earlier events receive higher predicted risk scores. Unlike RMSE or R^2, which require fully observed outcomes, or binary metrics that ignore time, the C-index respects the temporal ordering of events without requiring exact event times for all plants. For interpretability at individual survey dates, model performance was also summarized using standard binary classification metrics, including precision (the proportion of predicted critical plants that were truly critical), recall (the proportion of truly critical plants correctly identified), and their harmonic mean, the F1 score.

2.4.2. Post-Processing of RSF Risk Scores for Visualization

Continuous RSF-derived risk scores were normalized to a 0–1 scale within each flight to support visualization and temporal summaries. Because normalization was performed independently for each flight, risk classes represent relative risk within a given survey date rather than absolute risk comparable across dates. Temporal changes in disease progression were therefore evaluated using the underlying continuous RSF risk scores and time-to-event structure of the survival model, which preserve information on relative risk accumulation over time. Changes in the proportion and spatial extent of moderate- and high-risk plants across flights reflect the expansion of higher-risk zones rather than direct comparisons of absolute risk magnitude between survey dates. To facilitate visualization, normalized risk scores were discretized into three canopy risk classes based on empirical inspection of risk score distributions across flights. Plants with normalized risk ≤ 0.15 were classified as low risk, those with $0.15 <$ normalized risk ≤ 0.40 as moderate risk, and those with normalized risk > 0.40 as high risk. These categories were used solely for risk maps and percent-at-risk summaries; all statistical

 doi:10.20944/preprints202602.1030.v1

analyses and post-hoc treatment comparisons were conducted using continuous RSF risk scores while controlling for field and flight date.

2.4.3. Secondary Classification Analysis

To complement the time-to-event analysis, we also evaluated RSF performance as a per-flight binary classifier. For each test flight, continuous risk scores were thresholded using the value that maximized the F1 score, and plants in the cluster-defined critical group were treated as positive cases. Precision, recall, and F1 scores were computed using the *scikit-learn* package to quantify classification performance across flights.

2.5. Post-Hoc Influence of Variable Fumigation Treatment

Because fumigation rates were applied operationally across the study fields based on historical disease pressure, we conducted a post-hoc analysis to evaluate whether RSF-derived canopy risk varied systematically with fumigation intensity.

The study fields were part of a commercial strawberry operation, and pre-plant fumigation treatments were applied operationally rather than as a controlled experiment. Three chloropicrin–1,3-D (Pic-Clor 60) application rates were used across the site (standard = 392 kg/ha; medium = 336 kg/ha; low = 280 kg/ha), arranged in large contiguous blocks selected by the grower based on historical disease pressure and field logistics rather than randomized assignment. As a result, treatment areas were not replicated experimentally, and comparisons represent post-hoc observational patterns rather than causal contrasts.

Fumigation treatment polygons were delineated using field GPS surveys and digitized for spatial analysis. For each UAV survey date, plant-level hexagon units containing RSF-derived risk scores were spatially joined to the fumigation polygons based on spatial overlap. Hexagons intersecting low- or medium-rate polygons were labeled accordingly; all remaining hexagons were assigned to the standard rate category. A summary of fumigation-block areas and associated plant monitoring units within each rate class is provided in (Table 2).

Within-field treatment effects were evaluated using plant-level normalized RSF-derived canopy risk as the response variable. To address unequal sample sizes among rates, analyses were conducted on balanced subsamples created by random downsampling to equalize the number of plants per available rate within each field. Treatment effects were estimated using linear models with fumigation rate as a fixed effect and flight date included as a covariate, with standard errors clustered by flight date. As a robustness check, mixed-effects models with flight date specified as a random intercept were also fit to evaluate sensitivity of treatment effect estimates to model structure.

3. Results

3.1. Spatial Patterns of Decline

Among the three study fields, Field 1 exhibited the highest overall disease pressure, with clear expansion of canopy decline over the observation period (Figure 2). By June 1, early-season hotspots had coalesced into larger patches of moderate- and high-risk plants, particularly along central and northern regions. Time-series summaries reflected this trajectory: moderate-risk plants increased from 0.2% on May 15 to 13% on June 1, while high-risk plants rose from 0.87% to 1.96% over the same interval (Figure 2). By June 10, combined moderate- and high-risk plants comprised 26% of the field, indicating widespread late-season canopy collapse consistent with the mapped spatial patterns.

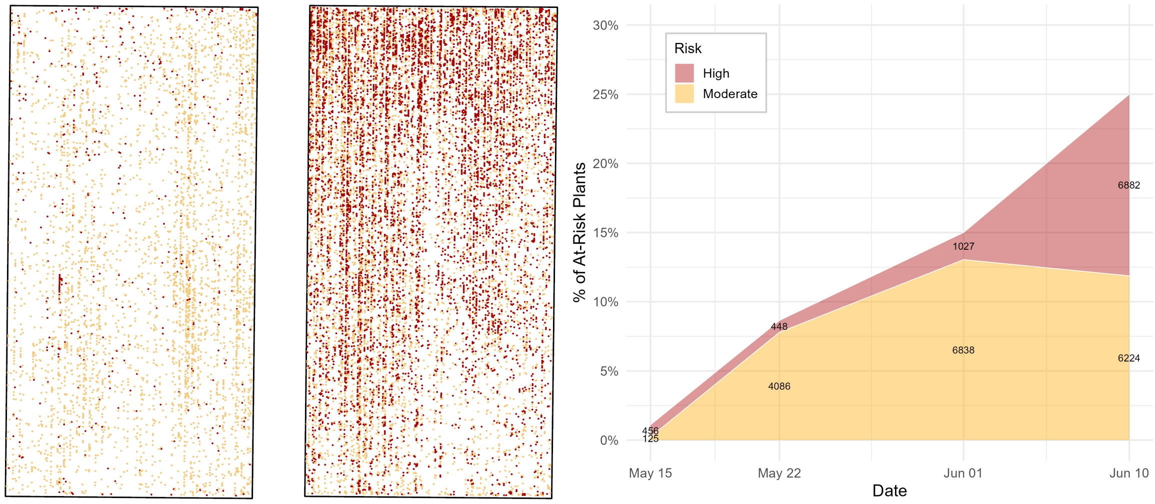

Figure 2. Spatiotemporal patterns of RSF-derived canopy risk in Field 1. Plant-level risk maps illustrate moderate and high canopy risk on 22 May 2023 (left) and 10 June 2023 (right), with the accompanying time series showing changes in the proportion of at-risk plants across survey dates.

In contrast, Field 2 showed consistently low and spatially diffuse canopy decline, with no evidence of sustained clustering. RSF-derived risk maps revealed scattered individual plants classified as moderate or high risk at different survey dates, with little temporal persistence of risk status at the individual-plant level and no indication of spatial expansion (Figure 3). Across the season, moderate-risk plants remained below 1% of the field (0.90% on May 15 and 0.28% on June 10), and high-risk plants remained below 3% (2.25% on May 15 and 2.02% on June 10), corresponding to isolated, low-intensity patches visible in the June imagery. Although overall decline incidence was low and individual risk status was often transient, the RSF consistently identified localized at-risk plants, indicating that subtle early stress signals were detectable even in a field with minimal visible symptoms.

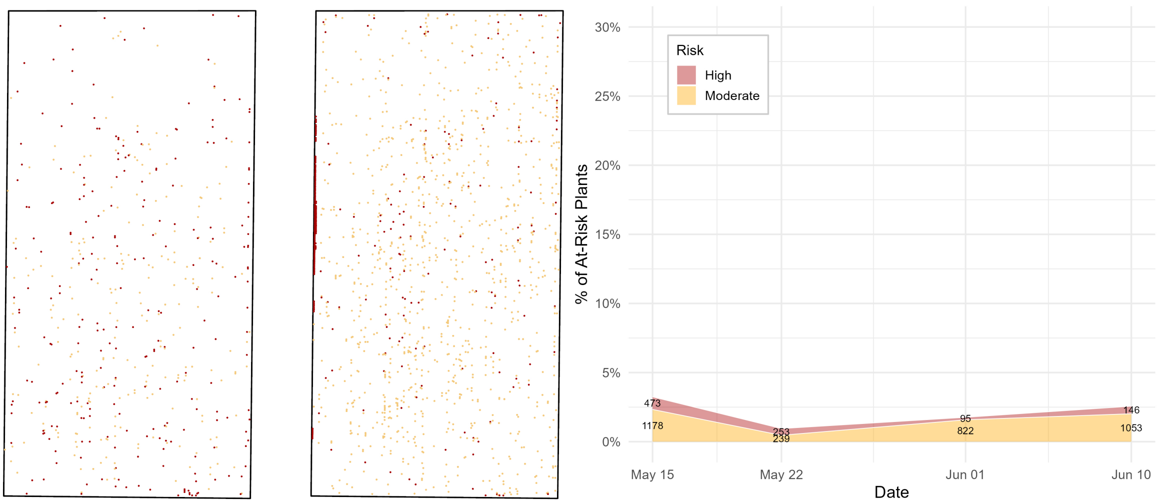

Figure 3. Spatiotemporal patterns of RSF-derived canopy risk in Field 2. Plant-level panels depict survey dates 22 May 2023 (left) and 10 June 2023 (right) as representative examples. Risk remained spatially diffuse with consistently low proportions of moderate- and high-risk plants across survey dates, as summarized in the accompanying time series.

Field 3 experienced limited field-wide decline, with risk concentrated primarily in two adjacent central rows (Figure 4). RSF-derived maps showed that most of the field remained uniformly healthy, while isolated moderate- and high-risk plants appeared early and recurred within the same localized spatial zones through June, as evidenced by consistent patch-level patterns across survey dates. Moderate-risk plants accounted for 4.22% of the field (2,300 plants) on May 15, a pattern consistent with

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 12 February 2026 doi:10.20944/preprints202602.1030.v1

transient physiological stress as plants entered peak fruit production, but declined to approximately 1% thereafter (0.43–1.17%). High-risk plants remained below 1% on all survey dates (0.14–0.95%). These patterns indicate limited field-wide decline but highlight localized zones of elevated stress that were not apparent from visual inspection alone. Even under low-symptom conditions, the RSF model consistently identified spatially coherent pockets of elevated risk in Field 3.

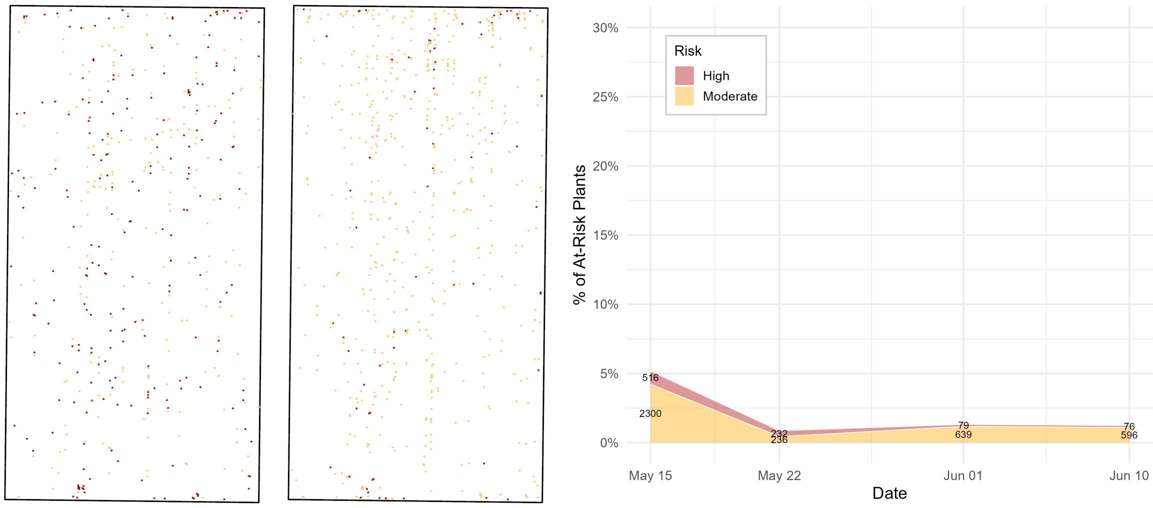

Figure 4. Spatiotemporal patterns of RSF-derived canopy risk in Field 3. Plant-level panels depict survey dates 22 May 2023 (left) and 10 June 2023 (right) as representative examples. Risk was concentrated in localized zones within the field, while the majority of plants remained low risk throughout the season, as reflected in the accompanying time series.

3.2. Random Survival Forest Model Performance

All reported performance metrics reflect out-of-sample evaluation on withheld plants and temporally forward test flights, with critical clusters defined using training data only, ensuring that model performance reflects generalization rather than memorization.

Model performance metrics including precision, recall, F1 score, and concordance (C-index) revealed strong field-to-field differences that reflected underlying canopy stress dynamics (Table 3). RSF performance varied across fields and survey dates in ways that corresponded to differences in the intensity and spatial coherence of canopy decline.

In Field 1, where symptoms were widespread and progressed predictably, models achieved high concordance throughout the season (C-index $\geq$ 0.97). Classification performance was strongest near peak decline on 1 June (F1 = 0.85), when spatial patterns were most coherent, but declined on 10–20 June (F1 = 0.18–0.09) as canopy collapse became visually saturated and class imbalance increased.

In Fields 2 and 3, where decline incidence was low, classification metrics were generally lower early in the season (F1 = 0.36–0.51 on 15–22 May) but improved during periods of greater spatial structure (for example, Field 3 with F1 = 0.84–0.83 on 10–20 June). Despite low prevalence, concordance remained consistently high across both fields (C-index $\geq$ 0.65), indicating that RSF models reliably ranked plants by relative risk even when absolute prediction accuracy was constrained by sparse events.

Taken together, these results show that RSF performance depends strongly on the spatial and temporal coherence of canopy decline. Classification accuracy peaks when symptoms are coordinated and widespread, whereas ranking performance as measured by the C-index remains robust even under minimal visible decline.

Importantly, RSF performance varied substantially across fields and survey dates, with strongest performance occurring during periods of spatially coherent canopy decline and reduced performance under low-prevalence conditions. This field- and time-specific variation indicates that model performance was driven by underlying canopy stress dynamics rather than by the modeling framework alone.

Table 3. Random Survival Forest (RSF) testing performance by field and survey date, summarized using precision, recall, F1 score, and concordance index (C-index).

	Field 1				Field 2				Field 3			
Flight Date	Precision	Recall	F1	C	Precision	Recall	F1	C	Precision	Recall	F1	C
4/21/2023	0.95	0.93	0.94	1.00	0.96	0.98	0.97	1.00	1.00	0.91	0.95	1.00
5/15/2023	0.67	0.39	0.49	0.68	0.29	0.59	0.39	0.83	1.00	0.35	0.51	0.86
5/22/2023	0.99	0.60	0.75	0.98	0.93	0.23	0.36	0.65	0.97	0.31	0.47	0.73
6/1/2023	0.89	0.81	0.85	0.99	0.86	0.87	0.87	0.96	0.86	0.82	0.84	0.89
6/10/2023	0.51	0.11	0.18	0.97	0.63	0.67	0.65	1.00	0.76	0.92	0.83	1.00
6/20/2023	0.38	0.05	0.09	0.68	0.68	0.33	0.44	1.00	0.87	0.79	0.83	1.00

3.3. Post-Hoc Comparison of Variable Fumigation Rates

After accounting for field- and flight-level effects, RSF-derived canopy risk did not differ meaningfully among fumigant application rates, consistent with the project objective of evaluating reduced fumigation rates without compromising plant performance. Within each field, normalized RSF risk scores were compared among available fumigation rates (280, 336, and 392 kg Pic-Clor 60/ha) using balanced subsamples and linear models that controlled for flight date (Table 4). Across all within-field contrasts, estimated treatment effects were very small (Δ normalized risk ≈ -0.002 to 0.023 on a 0–1 scale), and 95% confidence intervals from the primary, flight-clustered models overlapped zero (e.g., Field 1, 392 vs. 280 kg/ha: $\Delta = 0.022$, 95% CI -0.003 to 0.047; Field 3, 392 vs. 280 kg/ha: $\Delta \approx 0.000$, 95% CI -0.0018 to 0.0018). Empirical cumulative distribution functions showed nearly identical risk distributions among rates within fields. These differences were an order of magnitude smaller than field-level differences in canopy risk and are unlikely to be biologically meaningful, indicating no detectable association between pre-plant fumigation rate and RSF-estimated canopy risk in this observational setting.

Table 4. Within-field contrasts in normalized RSF-derived canopy risk among fumigant application rates, estimated using post-hoc linear models fit to balanced subsamples with flight-clustered standard errors.

Field	Treatment (kg/ha)	Effect Size (Δ Normalized Risk)	95% CI	n
1	392 vs. 280	0.022	-0.003, 0.047	49,224
2	392 vs. 336	−0.002	-0.009, 0.004	41,460
3	336 vs. 280	0.000	-0.001, 0.001	134,946
3	392 vs. 280	0.000	-0.002, 0.002	134,946

4. Discussion

This study demonstrates that UAV multispectral time series, consisting of repeated high-resolution surveys of the same plants across the growing season, combined with Random Survival Forest (RSF) modeling, can capture fine-scale spatiotemporal patterns of canopy decline in commercial strawberry systems. Across three fields spanning a gradient of disease pressure, RSF-derived risk surfaces reflected both the intensity and spatial organization of canopy stress, enabling early identification of decline under conditions ranging from widespread disease (Field 1) to low, spatially diffuse stress (Fields 2 and 3). These findings highlight the value of time-to-event modeling for cropping systems in which plant decline develops gradually and heterogeneously across space.

4.1. Capability Relative to Conventional Monitoring Approaches

Compared to conventional field scouting, single-date vegetation indices, or end-of-season yield assessments, the RSF-based framework provides a time-aware, plant-level view of canopy decline that captures both spatial structure and temporal progression.

The strongest model performance occurred in Field 1, where canopy decline was extensive and spatially coherent. In this field, early-season hotspots expanded into contiguous high-risk zones, and RSF models achieved consistently high concordance during periods of active decline (C-index ≥ 0.97), along with strong classification performance near peak disease expression (F1 = 0.85 on 1 June). These results indicate that survival-based models are particularly effective when stress progresses predictably and spectral signals accumulate over time. In contrast, Fields 2 and 3 exhibited limited visible decline, yet RSF predictions remained stable and consistently identified localized pockets of elevated risk. Although classification performance was reduced under these low-prevalence conditions (e.g., F1 = 0.36–0.51 early in the season), concordance remained high (C-index ≥ 0.83), indicating that the model preserved meaningful risk rankings even when binary prediction was constrained by sparse events.

Importantly, C-indices approaching 1.0 reflected accurate ranking of relative risk rather than overfitting. The concurrent decline in classification performance under low-incidence conditions supports this interpretation and is consistent with expected behavior of survival models applied to rare-event settings. Together, these results underscore a key insight: RSF performance depends not only on the amount of disease present but also on the spatial coherence of stress expression. When decline is widespread and coordinated across plants, predictive performance is high and risk maps align with visible symptoms. When stress is sparse or reflects transient physiological transitions, ranking performance remains informative even as classification accuracy decreases. This distinction is critical for operational use, where early-warning value may be derived from relative risk patterns before symptoms become visually obvious.

4.2. Value Added Beyond Vegetation Index Time Series

Trends in vegetation indices further support the biological relevance of the RSF predictions. Declines in NDRE and increases in Redness Index values are consistent with reduced chlorophyll content, canopy thinning, and early senescence. In the context of *Macrophomina phaseolina*, these spectral changes align with known symptomology of crown and root rot, which can impair water and nutrient uptake well before visible wilting occurs. Early-season moderate-risk signals observed in some fields, particularly Field 3, may reflect transient physiological stress during peak fruit production or early disease development that stabilizes later in the season.

While temporal trends in NDVI, NDRE, and RI reflected physiologically meaningful canopy stress, index trajectories alone do not account for censoring, nonlinear progression, or the spatial context of neighboring plants. In contrast, the RSF framework integrates these signals into a unified time-to-event model that explicitly handles plants that never decline, accommodates nonlinear interactions among predictors, and incorporates spatial proximity to stressed neighbors. This allows risk to be quantified even when spectral changes are subtle or transient, and explains why RSF-derived risk surfaces captured coherent spatial patterns in fields where index time series alone showed limited differentiation.

4.3. Applications and Transferability of the Framework

For soilborne pathogens that invade the root and vascular system, effective in-season chemical control options are limited once infection has occurred. However, early identification of high-risk zones may still inform stress-mitigating practices such as optimized irrigation management, which can alleviate water stress, delay symptom expression, or reduce plant mortality in some cases. In this context, the value of early detection lies less in curative intervention than in supporting informed, site-specific management decisions during the production season.

Spatial context emerged as a key component of model performance. Incorporation of proximity to critically stressed plants improved risk prediction, suggesting that local spatial clustering captures shared risk factors such as inoculum distribution, soil moisture gradients, or compaction patterns. The ability to detect localized stress zones, even in fields with minimal overall decline, highlights the advantage of plant-level spatial modeling over field-averaged metrics.

Post-hoc comparisons among variable fumigant application rates revealed no detectable association between fumigation intensity and RSF-estimated canopy risk in this observational layout. Effect sizes were small and confidence intervals overlapped zero, consistent with the non-randomized arrangement and lack of replication among treatment blocks. While these results do not support causal inference, they suggest that lower fumigation rates may be sufficient in portions of fields with historically low disease pressure, without clear increases in canopy risk during the growing season. This observation is consistent with the grower's operational decisions and highlights the potential for spatially differentiated fumigation strategies informed by historical disease patterns. In coastal California production systems, the ability to rotate away from infested fields is often constrained by land availability and cost, resulting in annual strawberry plantings into the same fields despite known pathogen presence. In this setting, UAV-derived risk maps provide actionable information for post-season management decisions, including fumigation planning, field selection, and input allocation for the subsequent production cycle.

Although this study evaluated post-hoc fumigation patterns as one applied example, the broader utility of this framework lies in its ability to support spatially targeted decision-making in high-value cropping systems. In strawberry production, this includes informing irrigation management, identifying zones for intensified scouting, prioritizing replant decisions, and guiding site-specific input allocation for subsequent seasons. Because the approach relies on general canopy stress responses rather than pathogen-specific signatures, it is readily transferable to other high-value crops where soilborne disease, water stress, or localized management constraints drive heterogeneous decline.

At the same time, the absence of replicated treatment blocks of comparable size and extent limits the statistical strength of this inference and points to important directions for future research. Replicated fumigation treatments across fields or seasons, with block sizes matched to management scales, would enable more rigorous evaluation of treatment effects. Integrating UAV-derived canopy risk maps with yield data and post-harvest assessments would further clarify whether localized reductions in fumigation intensity translate to comparable productivity outcomes. Such designs would allow formal testing of whether spatially targeted reductions in fumigant use can maintain yield while reducing chemical inputs.

A natural question is whether a single, pooled RSF model could be trained across multiple fields rather than fitting field-specific models. In this study, models were fit separately by field to account for differences in baseline disease pressure, soil conditions, and management history, which are known to influence canopy stress dynamics and survival risk. Although absolute risk levels varied among fields, the consistent performance of the RSF framework across contrasting disease regimes indicates that the underlying modeling approach is generalizable, even when parameterization is field-specific. Future work could explicitly evaluate pooled or hierarchical survival models that share information across fields while allowing field-level baseline hazards to vary, providing a pathway toward scalable deployment across production regions.

Several limitations should be acknowledged. Although the study fields have a documented history of *M. phaseolina*, pathogen presence was not confirmed for individual symptomatic plants, and spectral responses to disease may overlap with abiotic stressors such as nutrient limitation or irrigation-related water stress. The analysis was conducted over a single growing season and within a single pathogen context, and NDRE and RI thresholds were empirically defined rather than physiologically calibrated. Future work should focus on multi-season validation, integration of environmental covariates such as soil moisture and canopy temperature, and evaluation across additional cultivars and disease pressures. Further refinement of early-season thresholds may also improve pre-symptomatic detection.

Overall, this study highlights the potential of combining plant-level remote sensing with survival modeling to support early detection and spatial risk prediction of canopy decline in strawberry systems. By integrating temporal dynamics with spatial context, RSF models provide a promising framework for spatially informed, precision disease management in intensively managed cropping systems with recurring disease pressure.

Author Contributions: J.R.D. conceived and designed the analytical framework for the study, developed the modeling approach, performed the formal analysis, and prepared the original draft of the manuscript. C.A.G. conducted the UAV surveys, performed image processing and data preparation, and assisted with technical documentation for the manuscript. O.D. coordinated field site establishment, liaised with growers, and provided site-specific background information, and assisted with review of the manuscript. A.J.P. provided feedback on manuscript drafts and contributed to discussions on methodological development. F.S.M. contributed to the initial motivation and conceptual design for the study and assisted with review of the manuscript. F.N.M. developed and led the broader research project, provided supervision and critical review of the manuscript, and acquired funding. All authors have read and agreed to the published version of the manuscript.

Funding: This research was supported by the U.S. Department of Agriculture (USDA) through a grant from the Area Wide Pest Management Grant Program under Cooperative Agreement No. 58-2038-2-029. Contributions by F.S.M. were supported by the NASA Earth Action Agriculture Program Element.

Data Availability Statement: The data supporting the findings of this study, including UAV imagery and derived plant-level metrics, as well as the analysis scripts, are available from the corresponding author upon reasonable request due to data volume and grower confidentiality considerations.

Acknowledgments: The authors thank collaborating growers for providing access to field sites and logistical support during data collection, and acknowledge technical assistance provided by project staff during UAV data collection.

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

References

1. Simpson, D. The Economic Importance of Strawberry Crops. In *The Genomes of Rosaceous Berries and Their Wild Relatives*; Hytönen, T.; Graham, J.; Harrison, R., Eds.; Compendium of Plant Genomes, Springer: Cham, 2018. https://doi.org/10.1007/978-3-319-76020-9_1.
2. Food and Agriculture Organization of the United Nations. Crops and Livestock Products: Production Quantity. FAOSTAT Database, 2025. Data set; accessed 8 January 2026.
3. California Strawberry Commission. 2023 Acreage Survey. California Strawberry Commission, 2023. Accessed 8 January 2026.
4. Holmes, G.J. The California Strawberry Industry: Current Trends and Future Prospects. *International Journal of Fruit Science* **2024**, *24*, 115–129. https://doi.org/10.1080/15538362.2024.2342900.
5. California Department of Food and Agriculture. California Agricultural Production Statistics. California Department of Food and Agriculture, 2023. Accessed 8 January 2026.
6. Holmes, G.; Mansouripour, S.; Hewavitharana, S. Strawberries at the Crossroads: Management of Soilborne Diseases in California without Methyl Bromide. *Phytopathology* **2020**, *110*, 956–968. https://doi.org/10.1094/PHYTO-11-19-0406-IA.
7. Chamorro, M.; Miranda, L.; Domínguez, P.; Medina, J.J.; Soria, C.; Romero, F.; López-Aranda, J.M.; De los Santos, B. Evaluation of Biosolarization for the Control of Charcoal Rot Disease (Macrophomina phaseolina) in Strawberry. *Crop Protection* **2015**, *67*, 279–286. https://doi.org/10.1016/j.cropro.2014.10.021.
8. Alizadeh, M.; Khodadadi Manesh, S.; Fathi, P.; et al. Biology and Host Ranges of the Plant Pathogenic Fungus Macrophomina phaseolina: A Comprehensive Review. *Journal of Crop Health* **2025**, *77*, 50. https://doi.org/10.1007/s10343-024-01106-4.
9. Koike, S.T.; Arias, R.S.; Hogan, C.S.; Martin, F.N.; Gordon, T.R. Status of Macrophomina phaseolina on Strawberry in California and Preliminary Characterization of the Pathogen. *International Journal of Fruit Science* **2016**, *16*, 148–159. https://doi.org/10.1080/15538362.2016.1195313.
10. Steele, M.; Hewavitharana, S.; Henry, P.; Goldman, P.; Holmes, G. Survey of Late-Season Soilborne Pathogens Infecting Strawberry in Watsonville–Salinas, California. *Plant Health Progress* **2023**, *24*, 104–109. https://doi.org/10.1094/PHP-06-22-0056-S.
11. Steele, M.E.; Mendez, M.; Hewavitharana, S.S.; Holmes, G.J. Survey of Soilborne Pathogens Infecting Strawberry in Santa Maria, California. *International Journal of Fruit Science* **2024**, *23*, 256–266. https://doi.org/10.1080/15538362.2023.2292130.

12. Dilla-Ermita, C.J.; Goldman, P.H.; Jaime, J.; Ramos, G.; Pennerman, K.; Henry, P.M. First Report of Fusarium oxysporum f. sp. fragariae Race 2 Causing Fusarium Wilt of Strawberry (Fragaria × ananassa) in California. *Plant Disease* **2023**, *107*, 2849. https://doi.org/10.1094/PDIS-02-23-0217-PDN.
13. California Department of Pesticide Regulation. DPR Releases First Phase of State-Funded Fumigant Alternatives Study Focused on 1,3-D and Chloropicrin. Press release, 2025. Accessed 8 January 2026.
14. Nutter, F. W., J. The Role of Plant Disease Epidemiology in Developing Successful Integrated Disease Management Programs. In *General Concepts in Integrated Pest and Disease Management*; Ciancio, A.; Mukerji, K.G., Eds.; Springer, 2007; Vol. 1, pp. 203–226. https://doi.org/10.1007/978-1-4020-6061-8_3.
15. Oerke, E.C. Remote Sensing of Diseases. *Annual Review of Phytopathology* **2020**, *58*, 225–252. https://doi.org/10.1146/annurev-phyto-010820-012832.
16. Gogoi, N.K.; Deka, B.; Bora, L.C. Remote Sensing and Its Use in Detection and Monitoring of Plant Diseases: A Review. *Agricultural Reviews* **2018**, *39*, 307–313. https://doi.org/10.18805/ag.R-1835.
17. Mahlein, A.K.; Rumpf, T.; Welke, P.; Dehne, H.W.; Plümer, L.; Steiner, U.; Oerke, E.C. Development of Spectral Indices for Detecting and Identifying Plant Diseases. *Remote Sensing of Environment* **2013**, *128*, 21–30. https://doi.org/10.1016/j.rse.2012.09.019.
18. Calderón, R.; Navas-Cortés, J.A.; Lucena, C.; Zarco-Tejada, P.J. High-Resolution Airborne Hyperspectral and Thermal Imagery for Early Detection of Verticillium Wilt of Olive Using Fluorescence, Temperature and Narrow-Band Spectral Indices. *Remote Sensing of Environment* **2013**, *139*, 231–245. https://doi.org/10.1016/j.rse.2013.07.031.
19. Kouadio, L.; El Jarroudi, M.; Belabess, Z.; Laasli, S.E.; Roni, M.Z.K.; Amine, I.D.I.; Mokhtari, N.; Mokrini, F.; Junk, J.; Lahlali, R. A Review on UAV-Based Applications for Plant Disease Detection and Monitoring. *Remote Sensing* **2023**, *15*, 4273. https://doi.org/10.3390/rs15174273.
20. Chakhvashvili, E.; Machwitz, M.; Antala, M.; et al. Crop Stress Detection from UAVs: Best Practices and Lessons Learned for Exploiting Sensor Synergies. *Precision Agriculture* **2024**, *25*, 2614–2642. https://doi.org/10.1007/s11119-024-10168-3.
21. Mahlein, A.K. Plant Disease Detection by Imaging Sensors: Parallels and Specific Demands for Precision Agriculture and Plant Phenotyping. *Plant Disease* **2016**, *100*, 241–251. https://doi.org/10.1094/pdis-03-15-0340-fe.
22. Zarco-Tejada, P.J.; González-Dugo, M.V.; Fereres, E. Seasonal Stability of Chlorophyll Fluorescence Quantified from Airborne Hyperspectral Imagery as an Indicator of Net Photosynthesis in the Context of Precision Agriculture. *Remote Sensing of Environment* **2016**, *179*, 89–103. https://doi.org/https://doi.org/10.1016/j.rse.2016.03.024.
23. Wu, G.; Fang, Y.; Jiang, Q.; Cui, M.; Li, N.; Ou, Y.; et al. Early Identification of Strawberry Leaf Diseases Utilizing Hyperspectral Imaging Combined with Spectral Features, Multiple Vegetation Indices and Textural Features. *Computers and Electronics in Agriculture* **2023**, *204*, 107553. https://doi.org/10.1016/j.compag.2022.107553.
24. Rojas, A.E.F.; González, L.C.; Carrillo, D.A.P. Firma Espectral de Peca de la Hoja (Mycosphaerella fragariae (Tul.)) en Plantas de Fresa (Fragaria × ananassa Duch) Relacionada a NDVI y NDRE. *Ingeniería y Ciencia* **2023**, *19*, Article 4. https://doi.org/10.17981/ingecuc.19.2.2023.04.
25. Mahmud, M.S.; Zaman, Q.U.; Esau, T.J.; Chang, Y.K.; Price, G.W.; Prithiviraj, B. Real-Time Detection of Strawberry Powdery Mildew Disease Using a Mobile Machine Vision System. *Agronomy* **2020**, *10*, 1027. https://doi.org/10.3390/agronomy10071027.
26. Chang, Y.K.; Mahmud, M.S.; Shin, J.; Nguyen-Quang, T.; Price, G.W.; Prithiviraj, B. Comparison of Image Texture Based Supervised Learning Classifiers for Strawberry Powdery Mildew Detection. *AgriEngineering* **2019**, *1*, 434–452. https://doi.org/10.3390/agriengineering1030032.
27. Mahbub, J.; Bhandari, S.; Raheja, A.; Gonzalez, C. Machine Learning Models to Detect Strawberry Plant Health from UAVs for Real-Time Applications. In Proceedings of the Autonomous Air and Ground Sensing Systems for Agricultural Optimization and Phenotyping X. SPIE, 2025, Vol. 13475, pp. 213–226. https://doi.org/10.1117/12.3054123.
28. Bzdok, D.; Altman, N.; Krzywinski, M. Statistics versus Machine Learning. *Nature Methods* **2018**, *15*, 233–234. https://doi.org/doi.org/10.1038/nmeth.4642.
29. Zheng, C.; Abd-Elrahman, A.; Whitaker, V. Remote Sensing and Machine Learning in Crop Phenotyping and Management, with an Emphasis on Applications in Strawberry Farming. *Remote Sensing* **2021**, *13*, 531. https://doi.org/10.3390/rs13030531.

30. Ishwaran, H.; Kogalur, U.B.; Blackstone, E.H.; Lauer, M.S. Random Survival Forests. *The Annals of Applied Statistics* **2008**, *2*, 841–860. https://doi.org/10.48550/arXiv.0811.1645.
31. Rumpf, T.; Mahlein, A.K.; Steiner, U.; Oerke, E.C.; Dehne, H.W.; Plümer, L. Early Detection and Classification of Plant Diseases with Support Vector Machines Based on Hyperspectral Reflectance. *Computers and Electronics in Agriculture* **2010**, *74*, 91–99. https://doi.org/10.1016/j.compag.2010.06.009.
32. Kreigler, F.J.; Malila, W.A.; Nalepka, R.F.; Richardson, W. Preprocessing Transformations and Their Effects on Multispectral Recognition. In Proceedings of the Proceedings of the Sixth International Symposium on Remote Sensing of Environment, Ann Arbor, MI, USA, 1969; Vol. 1, pp. 97–131.
33. Gitelson, A.; Merzlyak, M.N. Quantitative Estimation of Chlorophyll-a Using Reflectance Spectra: Experiments with Autumn Chestnut and Maple Leaves. *Journal of Photochemistry and Photobiology* **1994**, *13*, 247–252. https://doi.org/10.1016/1011-1344(93)06963-4.
34. Hunt, E.R.J.; Cavigelli, M.; Daughtry, C.S.T.; McMurtrey III, J.E.; Walthall, C.L. Evaluation of Digital Photography from Model Aircraft for Remote Sensing of Crop Biomass and Nitrogen Status. *Precision Agriculture* **2005**, *6*, 359–378. https://doi.org/10.1007/s11119-005-2324-5.
35. Virtanen, P.; Gommers, R.; Oliphant, T.E.; et al. SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python. *Nature Methods* **2020**, *17*, 261–272. https://doi.org/10.1038/s41592-019-0686-2.
36. Pölsterl, S. scikit-survival: A Library for Time-to-Event Analysis Built on Top of scikit-learn. *Journal of Machine Learning Research* **2020**, *21*, 1–6.