

LETTER • OPEN ACCESS

Canopy-normalized UAV measurement of fungicide efficacy in wheat rust: revealing treatment-induced bias in disease observability

To cite this article: Moussa El Jarroudi *et al* 2026 *Environ. Res. Commun.* **8** 071008

View the [article online](#) for updates and enhancements.

You may also like

- [The prognostic value of skin temperature variability analysis in clinical settings: a systematic review of current evidence](#)
Eubi Chan, Angelica Blotto and Ali R Mani
- [Targeting in-stream habitat suitability with a multi-criteria mapping framework](#)
Corey Dawson, Gavin Scott, James Veres et al.
- [Assessment of blue carbon storage and driving mechanisms in temperate seagrass beds of the Changshan Archipelago, North Yellow Sea](#)
Kaiting Zhi, Jun Yang, Fengjiao Zhang et al.

Environmental Research Communications



LETTER

OPEN ACCESS

RECEIVED
10 April 2026

REVISED
24 June 2026

ACCEPTED FOR PUBLICATION
1 July 2026

PUBLISHED
17 July 2026

Original content from
this work may be used
under the terms of the
[Creative Commons
Attribution 4.0 licence](#).

Any further distribution
of this work must
maintain attribution to
the author(s) and the title
of the work, journal
citation and DOI.



Canopy-normalized UAV measurement of fungicide efficacy in wheat rust: revealing treatment-induced bias in disease observability

Moussa El Jarroudi¹ , Louis Kouadio² , Jonathan Peereman^{1,*} and Marco Beyer³

¹ University of Liège, Department of Environmental Sciences and Management, SPHERES Research Unit, Arlon, Belgium

² Africa Rice Center, Bouake, Ivory Coast

³ Luxembourg Institute of Science and Technology, 41 Rue du Brill, L-4422 Belvaux, Luxembourg

* Author to whom any correspondence should be addressed.

E-mail: jonathan.peereman@uliege.be, meljarroudi@uliege.be, A.Kouadio2@cgiar.org and marco.beyer@list.lu

Keywords: RGB imagery, disease monitoring, yellow rust, leaf rust, canopy cover

Supplementary material for this article is available [online](#)

Abstract

Unmanned aerial vehicle UAV-based RGB imaging is increasingly used to track crop diseases and assess fungicide performance. However, standard disease severity measures may miss an important factor: fungicide-induced changes in canopy structure that modulate disease observability. Here we analyse a multi-site dataset of wheat yellow and leaf rust, integrating UAV-derived canopy cover and disease severity, and rater-based disease severity assessments. Results indicate that between visual and UAV imagery-based disease severity was high but fundamentally non-stationary, governed by canopy cover and epidemic intensity. Measurement error followed structured, canopy-dependent regimes. The canopy-normalized disease area, a metric that harmonizes severity with structural and spatial context, revealed ranking shifts in fungicide efficacy that are not captured through traditional metrics. This suggests that severity-based comparisons may combine true disease suppression with treatment-related changes in how visible symptoms are. Framing plant disease quantification as a measurement challenge, rather than only a detection, may support a more accurate assessment of treatment effects. Modelling the influence of canopy structure on symptom expression could provide a scalable and robust basis for disease monitoring and decision-making in precision agriculture.

1. Introduction

Plant diseases remain a major constraint on global food security, with fungal pathogens alone accounting for substantial yield losses in major cereal systems worldwide (Savary *et al* 2019, Fones *et al* 2020, Singh *et al* 2025). Over the past two decades, unmanned aerial vehicle (UAV)-based remote sensing has emerged as a powerful tool for crop health monitoring, enabling non-destructive, spatially explicit assessments at high temporal and spatial resolution (Mahlein 2016, Chin *et al* 2023, Kouadio *et al* 2023, Manoj *et al* 2025). In plant pathology, most UAV studies focus on detecting or mapping disease symptoms and on correlating spectral indices with visually assessed disease severity once lesions are visible (Anam *et al* 2024, Mahlein *et al* 2024, Shafay *et al* 2025).

This symptom-oriented paradigm implicitly assumes that disease severity is directly observable and that spectral signals provide unbiased proxies of the underlying epidemic process. Yet visual disease ratings are known to be subjective, scale-dependent, and sensitive to canopy architecture and observer perception (Bock *et al* 2014, Deery and Jones 2021, Bock *et al* 2022). From a measurement-science perspective, both human raters and UAV sensors should be regarded as imperfect instruments observing a latent disease state rather than a direct ground truth. This distinction between latent disease processes and observed signals remains largely underexplored in UAV-based plant disease monitoring, despite its fundamental implications for data interpretation and model validity. A critical but largely overlooked issue is

that fungicides alter canopy structure by delaying senescence, preserving green leaf area, and modifying canopy density and architecture (Milne *et al* 2003, Robert *et al* 2008, Baccar *et al* 2011, Vidal *et al* 2017). These structural changes modify the optical environment in which disease is observed. Thus, identical levels of underlying disease burden can produce different spectral responses depending on canopy state, leading to systematic distortions in remotely sensed disease estimates.

Comparisons of fungicide efficacy based solely on observed severity may conflate true disease suppression with treatment-induced changes in disease observability. Whereas UAV imagery-based disease studies have largely focused on enhancing detection accuracy or optimizing spectral features (Chin *et al* 2023, Mahlein *et al* 2024, Manoj *et al* 2025), this study explores disease quantification from a measurement perspective. In this view, disease severity is treated not as a directly observable variable, but as a latent process shaped by canopy structure and observation conditions. This perspective enables the identification of non-stationary measurement regimes—conditions under which the relationship between disease severity and UAV-derived signals varies in strength, bias, or uncertainty as a function of canopy or epidemic state—and provides a framework for interpreting variability in UAV-based estimates. Within this framework, canopy structure emerges as a key state variable influencing disease observability. By shaping background effects, leaf overlap, illumination heterogeneity, and symptom visibility, canopy cover acts as a structural filter that modulates the link between spectral signals and disease severity. As a result, measurement error is better viewed as structured and dynamic, rather than purely random, and dependent on canopy conditions.

This study aimed to develop a treatment-aware UAV measurement framework for yellow rust and leaf rust epidemics in wheat, accounting for canopy-mediated observability. Specifically, this study aimed at (i) quantifying human–UAV agreement under contrasting canopy cover regimes, (ii) demonstrating that measurement error and bias are structured rather than random, and (iii) introducing a canopy-normalized disease metric that yields robust and transferable fungicide rankings across sites and cultivars. By challenging the assumption that disease severity is directly observable, this study introduces a canopy-aware measurement framework in which plant disease is treated as a latent process filtered by canopy structure and observation conditions. Rather than viewing UAV-derived estimates as direct proxies of severity, we interpret them as measurement outputs shaped by canopy-mediated observability. This reframes UAV-based disease assessment from a detection-oriented task to a measurement problem, distinguishing our approach from studies focused mainly on classification accuracy or spectral feature optimization.

2. Conceptual framework

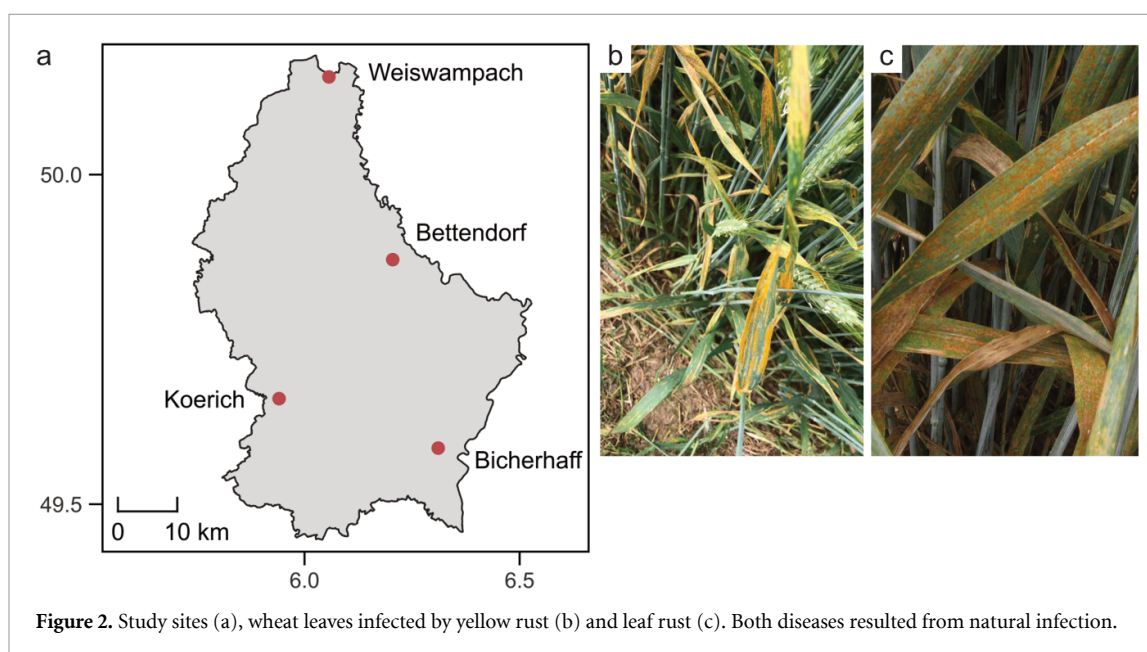
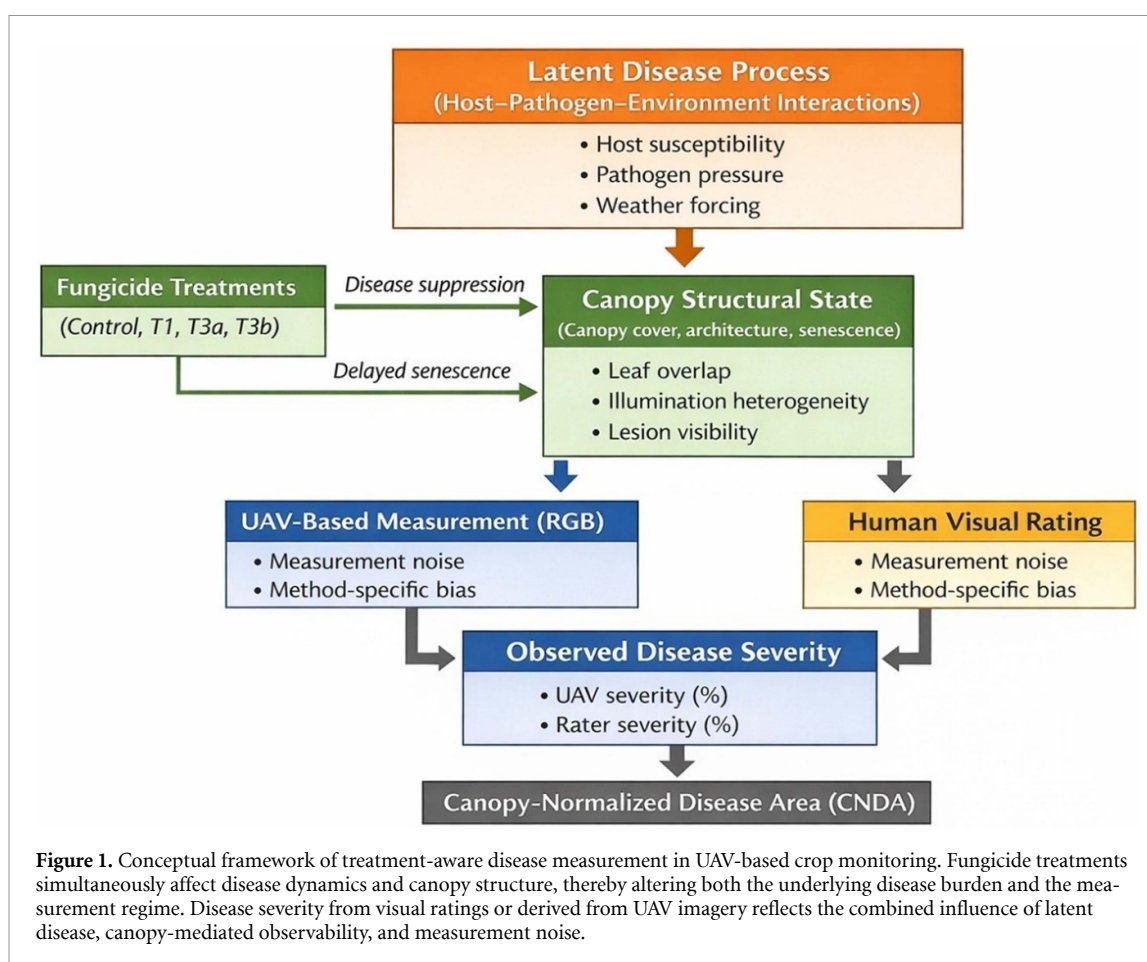
Disease development is a latent biological process governed by host–pathogen interactions and environmental forcing (Gilligan 2007). Observed disease severity represents a filtered manifestation of this process, shaped by canopy structure and the properties of the measurement instrument. Fractional canopy cover governs background effects, leaf overlap, illumination heterogeneity, and lesion visibility, thereby acting as a state variable controlling disease observability (figure 1). Within this framework, fungicide treatments influence disease dynamics and canopy structure simultaneously. As a result, observed severity reflects a joint process linking latent disease burden, canopy-mediated observability, and measurement noise. This perspective motivates agreement-based analyses and canopy-normalized metrics that explicitly account for treatment effects on the measurement regime.

3. Materials and methods

3.1. Experimental design and disease monitoring

Field trials were conducted across four representative winter wheat production sites in Luxembourg (Bettendorf, Bicherhaff, Koerich, and Weiswampach; figure 2) during the 2018–2019 growing season. The experimental design followed a fully randomized complete block structure with four replicates per treatment, each plot covering 12 m². Within each block, fungicide-treated and untreated control plots were randomly distributed to minimize spatial bias. Two winter wheat (*Triticum aestivum* L.) cultivars, Kerubino and Desamo, characterized by contrasting susceptibility levels to wheat yellow rust and wheat leaf rust (Bundessortenamt 2018), were selected.

These cultivars were subjected to four fungicide regimes: an untreated control, a single-application programme (T1), and multi-application programmes (T2/T3 equivalents, hereafter referred to as T3a and T3b depending on timing strategies) (table 1). Fungicide applications were synchronized with key



phenological stages (e.g. GS31–GS65), in accordance with standard agronomic recommendations and site-specific disease risk conditions. Crop management practices, including sowing dates, fertilization, and weed control, were consistent with regional farming practices to ensure agronomic realism.

Disease monitoring was conducted weekly from April to July 2019 through standardized visual assessments. For each plot, ten plants were randomly selected and permanently marked at the beginning of the season, resulting in 40 plants monitored per cultivar and per site. Disease incidence (proportion of infected plants) and severity (percentage of diseased leaf area) were assessed repeatedly on the same

Table 1. Overview of experimental design, fungicide treatments, and variables used in the study. GS31–GS65 indicate phenological stages for the fungicide treatments T3a and T3b. The fungicides were commercially available and applied according to manufacturer and official recommendations.

Category	Description
Sites	Bicherhaff, Bettendorf, Koerich, Weiswampach
Observation dates	15 May 2024May, 5 June, and 3 July 2019
Fungicide treatments	Control: no fungicide application T1: one application (1.5 l ha ^{−1} Swing Gold + 1 l ha ^{−1} Bravo) T3a: three applications (GS31: 0.75 l ha ^{−1} Opus Team + 1 l ha ^{−1} Bravo; GS37: 1.5 l ha ^{−1} Fandango + 1 l ha ^{−1} Bravo; GS59: 1.5 l ha ^{−1} Swing Gold + 1 l ha ^{−1} Bravo) T3b: three applications (GS31: 0.75 l ha ^{−1} Opus Team + 1 l ha ^{−1} Bravo; GS37: 1.5 l ha ^{−1} Fandango + 1 l ha ^{−1} Bravo; GS69: 1.5 l ha ^{−1} Swing Gold + 1 l ha ^{−1} Bravo)
UAV-derived variables	Fractional canopy cover (UAV_Fcover, %); UAV-based disease severity for yellow and leaf rust (%); affected disease area (m ²)
Visual assessments	Rater-based disease severity for yellow rust and leaf rust (%)
Derived metric	Canopy-normalized disease area (CNDA = Severity × UAV_Fcover × Plot area)
Total observations	320 plot-level observations

plants to capture temporal epidemic dynamics. This repeated-plant sampling strategy reduced within-plot variability and ensured that changes in disease incidence and severity reflected epidemic progression rather than differences in plant selection over time. All disease assessments explicitly distinguished between wheat yellow rust and leaf rust, and other prevailing diseases (*i.e.* Septoria leaf blotch, powdery mildew).

Visual assessments were conducted by four different raters: two plant pathologists with extensive experience in disease assessment and two agronomists also experienced in wheat disease assessment. To maintain consistency and minimize observer bias across assessment dates, the same rater assessed each replicate throughout the monitoring period. Stripe rust and leaf rust were evaluated separately throughout the monitoring period to preserve disease-specific epidemic patterns and avoid merging the dynamics of two pathosystems. Disease severity was visually assessed using the modified Cobb scale (Peterson *et al* 1948). To evaluate the reproducibility and consistency of visual disease assessments, inter-rater agreement was assessed using an independently scored subset of observations spanning the full range of canopy conditions, disease severities, cultivars, sites, and assessment dates encountered during the study. Agreement was quantified using concordance-based metrics, thereby allowing the evaluation of both measurement consistency and potential observer-dependent bias. An overview of experimental design, study sites, cultivars, fungicide treatments, observation dates, and measured variables is provided in table 1.

3.2. UAV data acquisition, canopy cover and disease severity estimation

High-resolution RGB imagery was acquired over all experimental plots at four key dates using an UAV equipped with an onboard RGB sensor (DJI Phantom 4 Pro, DJI, China), following a standardized acquisition protocol (Heidarian Dehkordi *et al* 2020). Flights were conducted at 10 m above ground level under nadir view, with an airspeed of 4 m s^{−1}, 85% front overlap, and 75% side overlap, resulting in a ground sampling distance of approximately 0.28 cm; flight missions were planned using Pix4Dcapture software (Pix4D S.A., Lausanne, Switzerland). To minimize radiometric inconsistencies across dates and sites, image acquisition was restricted to stable illumination conditions and preprocessing included radiometric normalization procedures. Image acquisition campaigns were synchronized with key phenological stages (from stem elongation to grain filling) and aligned with disease assessment dates to ensure temporal consistency between spectral and visual observations.

All images were processed using photogrammetric workflows (Pix4Dmapper) to generate georeferenced orthomosaics suitable for quantitative analysis of canopy cover and disease severity. Fractional canopy cover (UAV_Fcover, %) was derived from RGB imagery using vegetation segmentation techniques. In contrast to conventional approaches where canopy cover is treated as a secondary descriptor, UAV_Fcover was here explicitly considered as a state variable controlling disease observability, as it directly influences background reflectance, leaf overlap, illumination heterogeneity, and lesion visibility.

RGB orthomosaics were first radiometrically normalized to reduce illumination differences among acquisition dates and sites. Individual experimental plots were subsequently extracted from the orthomosaics using predefined plot boundaries. Vegetation pixels were segmented from soil and background elements using canopy-specific RGB thresholding procedures adapted to canopy development stage. Disease severity estimation was then performed through symptom extraction algorithms previously validated for wheat rust epidemics (Heidarian Dehkordi *et al* 2020), which identify rust-related discoloration patterns within segmented vegetation areas and quantify the proportion of symptomatic canopy pixels relative to total vegetation pixels.

The disease-severity workflow therefore consisted of four successive steps: (i) radiometric normalization of RGB orthomosaics, (ii) plot extraction, (iii) vegetation segmentation, and (iv) disease symptom extraction and severity quantification.

Quality-control procedures were applied throughout the processing chain and included visual inspection of orthomosaics, removal of image artefacts caused by motion blur or illumination anomalies, verification of plot boundaries, exclusion of incomplete plot images, and consistency checks among acquisition dates. These procedures ensured that disease estimates reflected biological variability rather than image-processing artefacts.

Disease severity for yellow rust and leaf rust was estimated from UAV RGB imagery using previously validated approaches (Heidarian Dehkordi *et al* 2020). It is important to note that both UAV-derived and visually assessed disease severities were treated as observational estimates of an underlying latent disease process, rather than as direct ground truth measurements, in line with measurement theory considerations (Bland and Altman 1986, Lin 1989).

3.3. The canopy-normalized disease area (CNDA)

To integrate disease intensity, the observable canopy domain, and the spatial extent over which symptoms occur, we defined a CNDA (equation (1)):

$$\text{CNDA} = \text{Severity} \times (\text{UAV_Fcover}/100) \times \text{Plot_area} \quad (1)$$

where Severity represents UAV-derived disease severity expressed as the percentage of diseased leaf area, UAV_Fcover represents UAV-derived fractional green canopy cover, and Plot_area corresponds to the surface area of each experimental plot.

The quantity $\text{Severity} \times \text{UAV_Fcover} \times \text{Plot_area}$ should be interpreted as an estimate of the effective observable diseased canopy area rather than the absolute area of infected leaf tissue. This distinction is central to the rationale of the metric. Disease severity alone provides information on the proportion of symptomatic tissue but does not account for the structural domain within which symptoms can develop and be detected. Consequently, two plots exhibiting identical severity values may differ substantially in the amount of observable disease if their canopy architectures differ. By incorporating UAV-derived canopy cover, CNDA accounts for these structural differences and provides a more realistic representation of disease expression at the canopy scale. The metric was therefore designed to capture the interaction between symptom intensity and canopy development, which is particularly relevant in UAV-based phenotyping where disease observability depends not only on symptom severity but also on the extent of the canopy available for symptom expression and detection.

For all calculations, disease severity and UAV_Fcover were converted from percentages to fractional values before multiplication. Thus, a plot exhibiting 10% disease severity and 80% canopy cover over a 12 m² plot corresponds to $\text{CNDA} = 0.10 \times 0.80 \times 12 = 0.96$ m²-equivalent. The resulting unit therefore represents an effective observable diseased canopy area rather than the absolute area of infected tissue. The CNDA metric was introduced to reduce treatment-induced bias in disease estimation because fungicide applications may modify canopy structure through delayed senescence, preservation of green leaf area, and changes in canopy density and architecture. These structural modifications can alter disease observability independently of the underlying disease process (Robert *et al* 2008, Vidal *et al* 2017).

Importantly, CNDA should not be interpreted as a direct estimate of pathogen biomass, infection intensity, or total biological disease burden. Rather, it quantifies disease expression within the observable canopy domain available to both visual raters and UAV sensors. Consequently, low CNDA values may arise either from low disease expression or from a reduced observable canopy domain caused by senescence, canopy degradation, or severe disease-induced tissue loss.

CNDA should therefore be interpreted as a canopy-aware observational metric designed to improve comparability among treatments exposed to contrasting canopy structural conditions rather than as a biologically absolute measure of disease burden. Its primary purpose is to standardize disease observations across heterogeneous canopy states and to reduce canopy-mediated distortions that may influence disease quantification.

3.4. Statistical analysis

Treatment effects were analysed using hierarchical mixed-effects models in R (version 4.4.1) (R Core Team 2024). Treatment, cultivar, and disease type (yellow rust or leaf rust) were modelled as fixed effects, while site, plot, and observation date were included as random effects to account for spatial variation, repeated measurements, and temporal dependence. Plot identity was retained to avoid pseudo-replication.

Complete mixed-effects model outputs are reported in supplementary table S1 to facilitate reproducibility and independent assessment of model performance. The supplementary information includes fixed-effect estimates, standard errors, confidence intervals, variance components, goodness-of-fit measures, residual diagnostics, and model validation statistics.

Rolling moving-window correlations were calculated using a five-point moving window. Boundary windows with fewer than five observations were excluded, and correlations were therefore computed only for complete windows. Correlations were calculated separately for yellow rust and leaf rust by pooling observations across the four experimental sites. Temporal trajectories of UAV-derived disease severity, visual disease severity, UAV_Fcover, and CNDA were analysed across observation dates, with mean profiles calculated separately for yellow rust and leaf rust. Following epidemiological frameworks (Jeger *et al* 2007), canopy structure and disease severity were treated as interacting processes.

Error landscapes were generated by pooling observations across the four sites to highlight generalizable patterns. Uncertainty in CNDA-based treatment rankings was assessed with a non-parametric bootstrap using 2000 resamples, from which treatment values, ranking positions, and 95% confidence intervals were recalculated. Agreement metrics (CCC, MAE, RMSE, and bias) were interpreted alongside canopy-class sample sizes and regime-consistency analyses, either by disease or jointly to identify broader patterns.

Agreement between UAV-derived and visually assessed disease severity was evaluated using Lin's concordance correlation coefficient (CCC) (Lin 1989), Pearson correlation coefficient, mean absolute error (MAE), root mean square error (RMSE), and Bland–Altman bias (Bland and Altman 1986). Together, these metrics describe association, agreement, and systematic deviation. CCC ranges from 0 (no concordance) to 1 (perfect concordance). Metrics were stratified according to canopy cover classes to identify canopy-dependent measurement regimes and evaluate whether UAV–visual agreement remained stable across canopy development stages.

Canopy cover was classified into three biologically meaningful categories based on canopy closure status and expected changes in symptom observability: low canopy cover (<60%), corresponding to open canopies with substantial soil background exposure; intermediate canopy cover (60%–80%), representing transitional canopy closure conditions; and high canopy cover (>80%), corresponding to near-complete canopy closure characterized by increased leaf overlap, reduced background effects, and altered lesion visibility. These thresholds were selected *a priori* to represent distinct structural states of canopy development rather than statistical quantiles of the dataset and are consistent with agronomic interpretations of crop canopy closure dynamics.

4. Results

4.1. Overall agreement between UAV and rater assessments

Inter-rater agreement remained consistently high across sites, cultivars, disease severity levels, and observation dates, indicating strong reproducibility of visual disease assessments and limited observer-dependent variability. These findings support the reliability of the visual reference dataset used throughout the study and reinforce confidence in subsequent UAV–rater agreement analyses.

Agreement between UAV- and rater-based estimates was influenced by canopy conditions. Indeed, although UAV- and rater-based disease severity estimates showed strong overall agreement for both yellow rust and leaf rust (figures 3(a) and (c)), dispersion increased with disease intensity. Lin's CCC confirmed high agreement overall but revealed substantial variability across canopy cover classes (table 2). The CCC decreased markedly with increasing canopy cover for both diseases, with values ranging from 0.883 to 0.667 for yellow rust, and 0.950–0.226 for leaf rust. The agreement metrics indicated a non-stationary human–UAV relationship, with concordance and bias strongly modulated by canopy cover.

4.2. Canopy-dependent measurement bias

UAV and visual assessments were closely aligned for both diseases. For yellow rust, UAV and rater estimates were strongly correlated (Pearson $r = 0.94$) (figure 3(a)). Similarly, for leaf rust, UAV and visual assessments were also strongly correlated (Pearson $r = 0.95$) (figure 3(a)). Although UAV-derived and

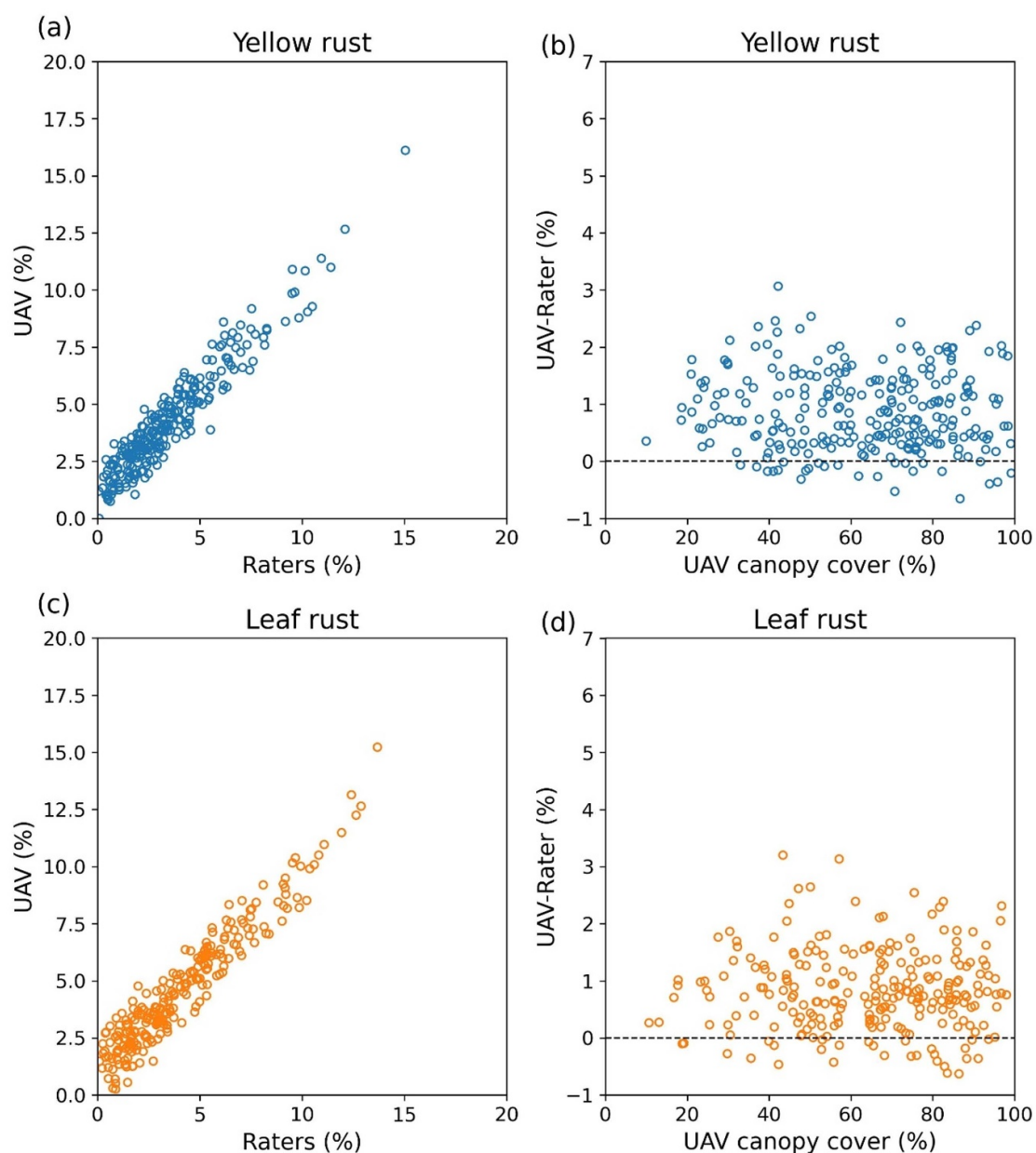


Figure 3. Comparison of UAV-derived and visual disease assessments for rust severity. Relationships between UAV-estimated severity and field raters' scores for yellow rust (a) and leaf rust (c). Absolute residuals ($|UAV-Rater|$) plotted against UAV canopy cover for yellow rust (b) and leaf rust (d). Dashed lines indicate zero bias. Variability increases at lower canopy cover, while a slight negative bias is observed at higher canopy cover, suggesting reduced sensitivity of UAV estimates under dense canopies.

Table 2. Agreement between UAV- and rater-based disease assessments across canopy cover classes. CCC: Lin's concordance correlation coefficient; MAE: mean absolute error; RMSE: root mean square error; Bias: systematic bias derived from Bland–Altman analysis. Canopy cover was grouped into three classes: low (<60%), medium (60%–80%), and high (>80%).

Disease	Canopy cover class	<i>n</i>	CCC (Lin)	MAE	RMSE	Bias (UAV–Rater)
Yellow rust	Low	102	0.883	1.297	1.752	0.319
	Medium	126	0.823	0.854	1.225	0.112
	High	44	0.667	0.814	1.064	−0.185
Leaf rust	Low	102	0.95	2.027	3.089	0.561
	Medium	126	0.681	0.84	1.287	0.378
	High	44	0.226	1.064	1.516	−0.801

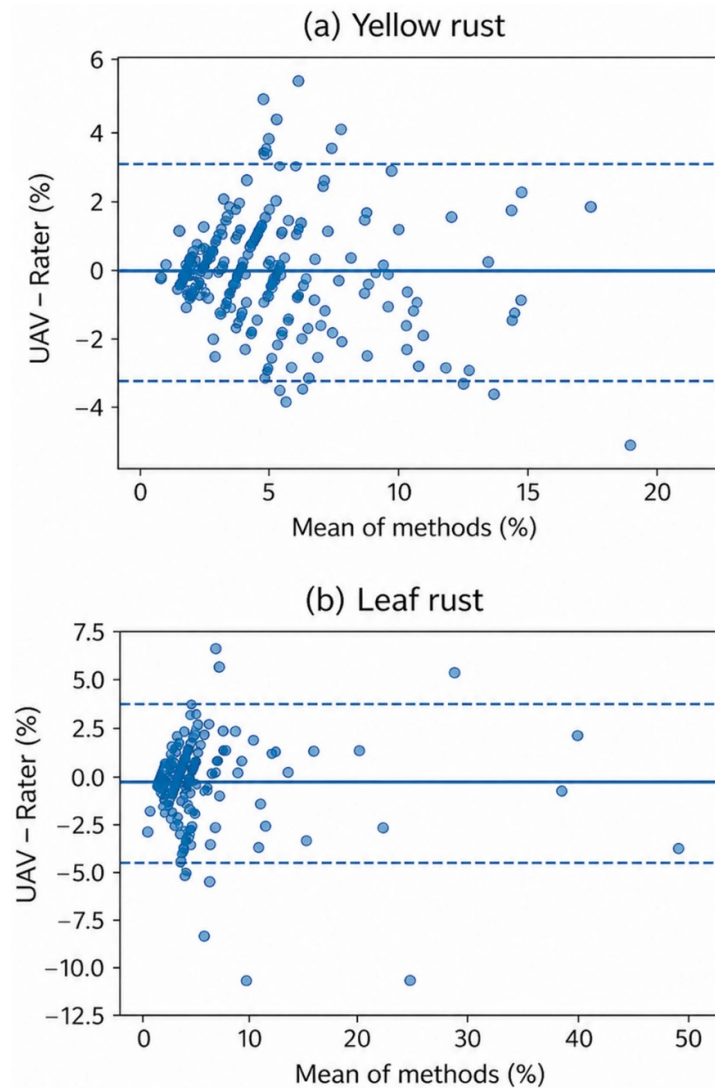


Figure 4. Bland–Altman plots showing the agreement between UAV-derived and rater-based estimates of yellow rust (a) and leaf rust (b) severity. The x-axis represents the mean disease severity obtained from UAV and visual assessments, while the y-axis represents the difference between methods (UAV–Rater). Solid horizontal lines indicate the mean bias, and dashed lines represent the 95% limits of agreement.

rater-based severity estimates were strongly correlated for both yellow rust and leaf rust (figure 3), this association did not indicate consistent measurement agreement. Instead, strong correlations coincided with substantial residual dispersion and pronounced canopy-dependent bias, particularly in residuals plotted against UAV-derived canopy cover (figures 3(b) and (d)). Differences between UAV and rater estimates decreased markedly as canopy cover increased (Pearson $r = -0.62$; figures 3(b) and (d)). Sparse canopy conditions were associated with larger and more variable errors, whereas near-complete canopy closure produced a stable measurement regime. Measurement error decreased significantly ($p < 0.001$) as canopy cover approached full closure, revealing a non-random, canopy-mediated regime of disease observability. Fungicide-induced canopy differences systematically affect disease measurement across site and cultivar. These results indicate that measurement error is structured rather than stochastic, following a canopy architecture–driven gradient with distinct regimes across the canopy closure continuum.

Bland–Altman analysis further highlighted systematic bias and wide limits of agreement under low canopy cover in both diseases (figure 4). Agreement for yellow rust remained stable across most severity ranges, with greater dispersion only at intermediate levels. Leaf rust showed much stronger heteroscedasticity and wider limits of agreement, especially in dense canopies and at higher severity.

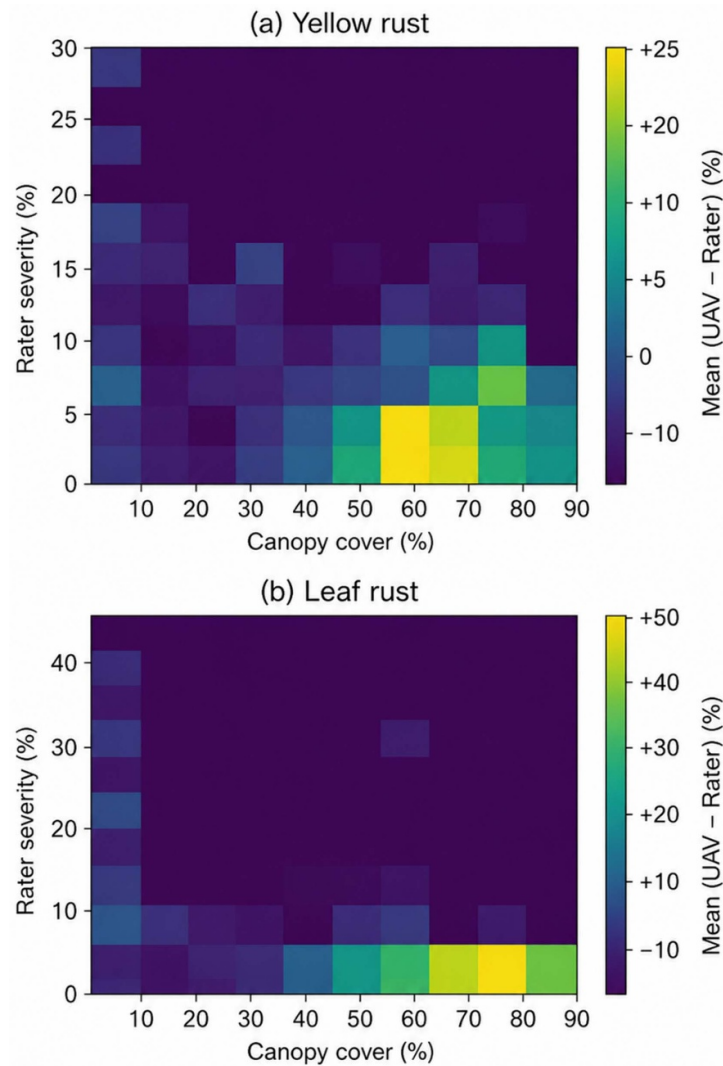


Figure 5. Aggregated error landscape of UAV-based disease estimation across canopy cover and disease severity gradients in winter wheat. (a) yellow rust; (b) leaf rust. Values were averaged across all experimental plots from the four study sites. Each cell corresponds to the mean error within a given canopy cover–disease severity bin.

4.3. Structured error across canopy cover and epidemic intensity

Mapping measurement error across the joint space of canopy cover and disease severity revealed a structured, non-random error landscape (figure 5). Error landscapes revealed structured patterns across canopy cover and disease severity gradients. These patterns indicate that canopy structure and disease severity cannot be treated as independent sources of variation but instead form a coupled system that jointly determines measurement error. High error values were consistently observed under sparse canopy conditions and during specific epidemic phases, with peak MAEs reaching approximately 29% for yellow rust and 50% for leaf rust. These patterns demonstrate that observability was jointly controlled by canopy structure and epidemic phase. Error patterns were not uniformly distributed but concentrated in specific regions of the canopy cover–disease severity space. This joint dependency confirms that both structural (canopy) and epidemiological (disease progression) processes interact to determine the reliability of UAV-based disease estimates.

The temporal analyses (figures 6 and 7) support this interpretation by showing that these error structures emerged and dissipated through time rather than remaining fixed throughout the season. Disease severity, canopy cover, and CNDA followed partially asynchronous trajectories (figure 6). The CNDA followed more regular temporal trajectories than raw severity, particularly during the transition between intermediate and high canopy cover stages (figure 6(c)), suggesting that canopy normalization not only improves spatial and treatment comparability but also reduces time-dependent distortions in disease measurement. Divergence between UAV-derived and visually assessed disease severity increased during

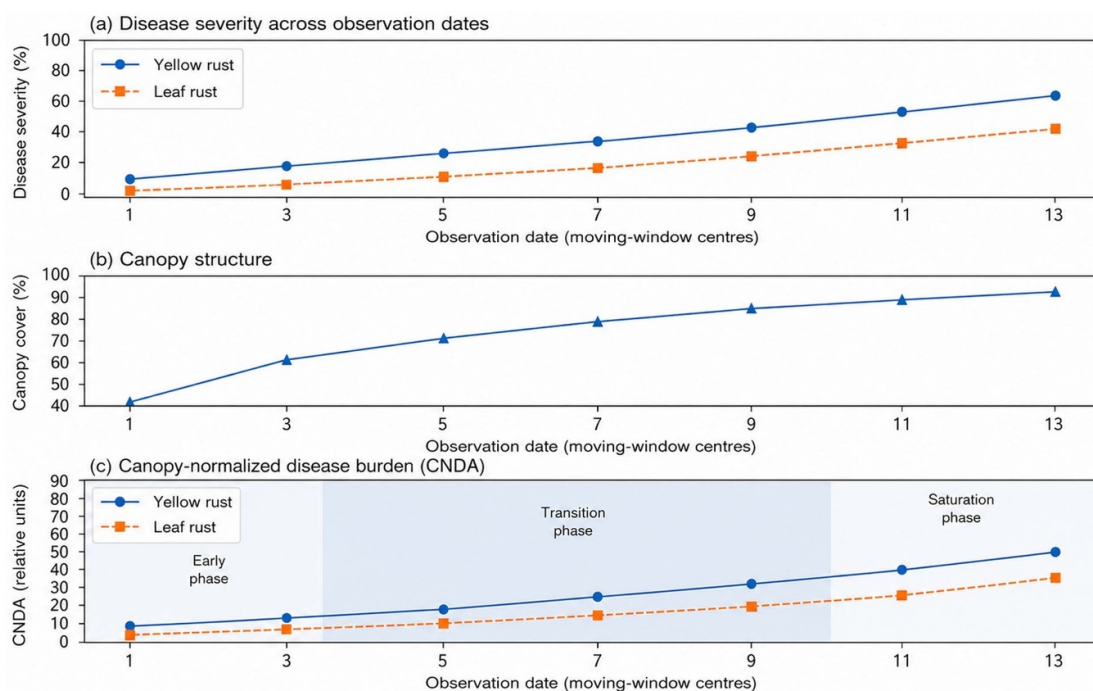


Figure 6. Temporal dynamics of disease severity (a), canopy structure (b), and canopy-normalized disease burden (CNDA) (c) across observation dates (moving window centre). Shaded regions in panel (c) represent distinct observational regimes corresponding to early, transition, and saturation phases of epidemic development. Solid blue and dashed orange lines in panels (a) and (c) represent yellow rust and leaf rust, respectively. Points on the x-axis correspond to interpolated moving-window centres.

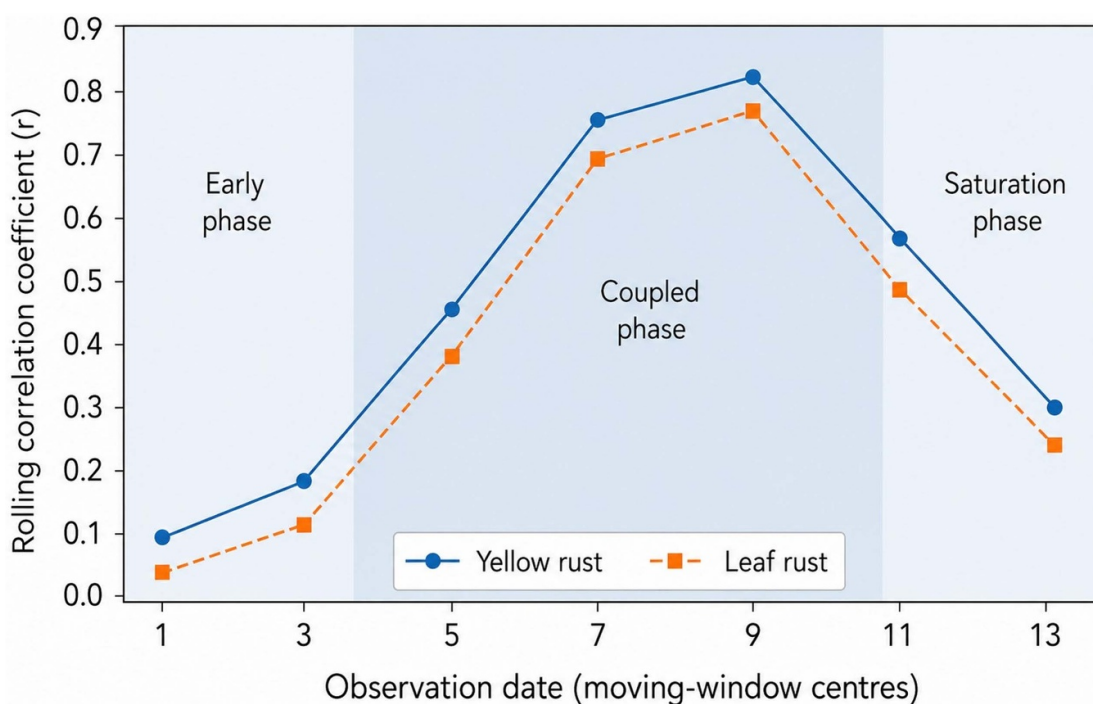


Figure 7. Temporal evolution of spectral-disease coupling across observation dates. Rolling moving-window correlation coefficients between UAV-derived disease estimates and visually assessed severity are shown for yellow rust (solid blue line with circles) and leaf rust (dashed orange line with squares). The x-axis represents moving-window centres derived from the interpolated temporal sequence. The light-blue, blue, and pale-blue shaded areas indicate the early, coupled, and saturation phases, respectively.

Table 3. Treatment effects on canopy-normalized disease burden with bootstrap confidence intervals. CI_low and CI_high correspond to the lower and upper bounds of the confidence intervals. T1 is a single application, T3a and T3b use three applications with different timings.

Disease	Treatment	Canopy-normalized disease area			
		Mean	CI_low	CI_high	Δ vs Control
Yellow rust	Control	0.0786	0.0613	0.0971	—
	T1	0.0854	0.0658	0.1071	0.0069
	T3a	0.0819	0.0628	0.1003	0.0033
	T3b	0.0982	0.0799	0.1184	0.0196
Leaf rust	Control	0.0429	0.032	0.0542	—
	T1	0.068	0.0469	0.097	0.0252
	T3a	0.0604	0.0464	0.0749	0.0175
	T3b	0.0515	0.0368	0.0675	0.0086

intermediate canopy stages, when canopy closure and epidemic intensity evolved simultaneously, and decreased again at later stages when canopy structure became more stable (figure 7).

The rolling correlation analysis (figure 7) provides an explicitly temporal expression of these regimes. In the early phase, correlation coefficients remained low and unstable, reflecting weak coupling between spectral response and disease severity. During the coupled phase, correlation increased markedly for both yellow rust and leaf rust, indicating optimal alignment between spectral signal and disease expression. In the saturation phase, correlation declined despite increasing disease severity, suggesting reduced spectral sensitivity under dense canopy conditions.

4.4. Treatment-induced confounding of severity-based comparisons

Fungicide treatments significantly modified canopy cover trajectories, particularly during later epidemic stages. Treatment rankings based on raw severity metrics varied across sites and dates, reflecting confounding between disease suppression and canopy-mediated observability. Fungicide treatments might have affected simultaneously both the biological process (disease progression) and the observation process (disease visibility), resulting in a treatment-induced confounding effect. Severity-based comparisons alone might misrepresent fungicide efficacy by capturing differences in observability rather than true differences in disease burden.

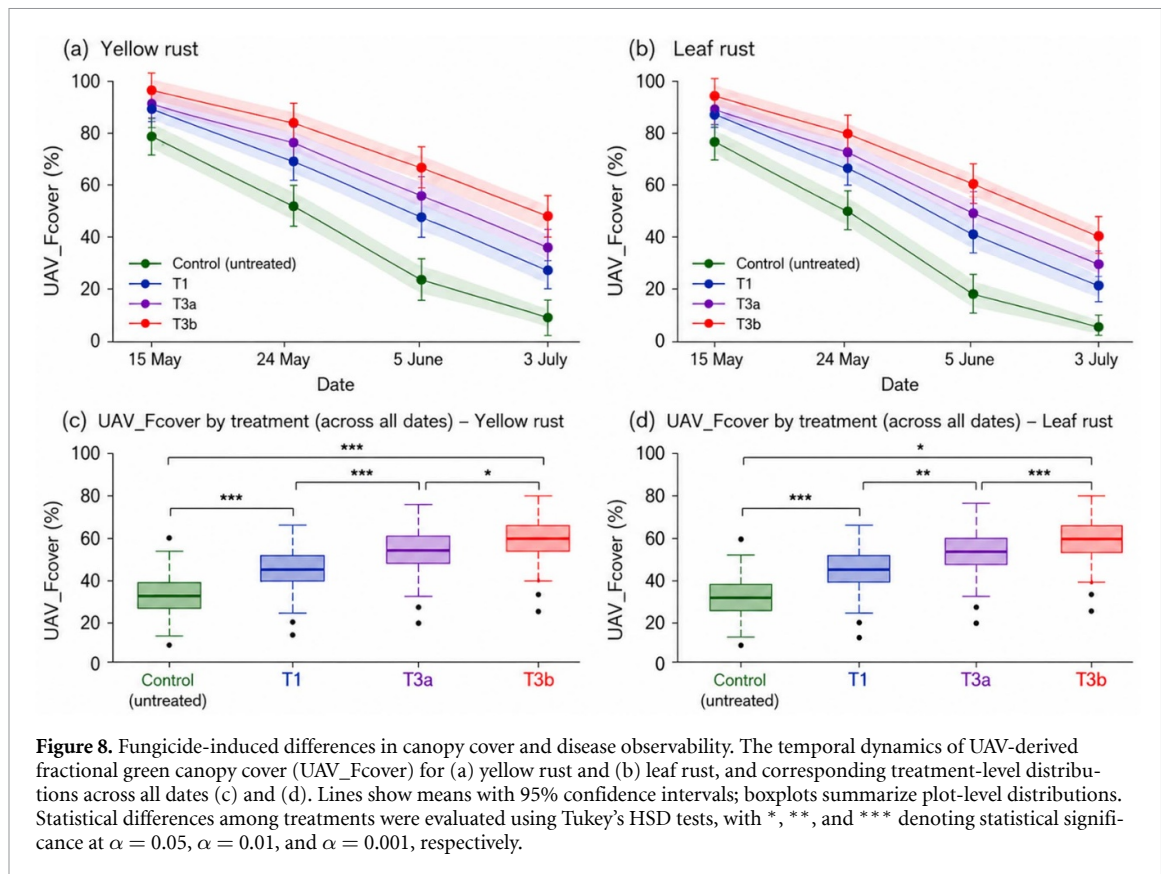
4.5. Canopy-normalized disease burden and treatment ranking

Table 3 presents the treatment effects on CNDA for both diseases. Confidence intervals associated with CNDA values are explicitly reported, providing a direct quantification of uncertainty. CNDA values were lower compared to those of fungicide-treated plots. This counterintuitive result does not indicate improved disease control but reflects a measurement artefact driven by canopy-mediated observability. Treated plots maintained denser canopies, increasing symptom visibility, whereas sparse canopies in untreated plots reduced lesion detectability. To further explore this potential mechanism, additional analyses were conducted using UAV-derived fractional canopy cover (UAV_Fcover) across fungicide treatments (figure 8). During the later stages of epidemic development, untreated plots consistently exhibited lower canopy cover than fungicide-treated plots, indicating that disease management influenced not only disease severity but also canopy structure (figures 8(a) and (b)). These differences in canopy architecture may, in turn, have affected symptom observability and therefore contributed to the observed CNDA trajectories. By maintaining larger fractions of healthy green canopy, fungicide-treated plots likely preserved a greater observable domain within which disease symptoms could be expressed and detected.

However, the overlap among bootstrap confidence intervals (table 3) indicates that CNDA should not be interpreted as providing statistically distinct separation among fungicide treatments. Rather, the metric primarily supports relative ranking consistency and effect-size interpretation while explicitly accounting for canopy-mediated observational variability.

The principal contribution of CNDA is therefore not to provide a biologically superior measure of disease burden, nor to replace conventional disease severity assessments, but rather to reduce canopy-mediated observational distortions that may influence treatment comparisons. By accounting for structural differences among canopies, CNDA facilitates more robust comparisons of disease expression across heterogeneous observation conditions.

Consequently, part of the treatment-related variation in CNDA may reflect differences in canopy structure and symptom visibility, in addition to differences in disease intensity alone (figures 8(c) and (d)).



Treatment rankings based on CNDA differed markedly from those derived from raw severity (figure 9). Canopy-normalized disease burden yielded a more discriminative and robust ranking of fungicide treatments than raw severity metrics. The CNDA reduced variability across sites and improved the consistency of treatment ranking, demonstrating its capacity to harmonize comparisons across heterogeneous canopy conditions. This confirms that integrating canopy structure into disease metrics is essential to obtain biologically meaningful and comparable estimates of treatment performance.

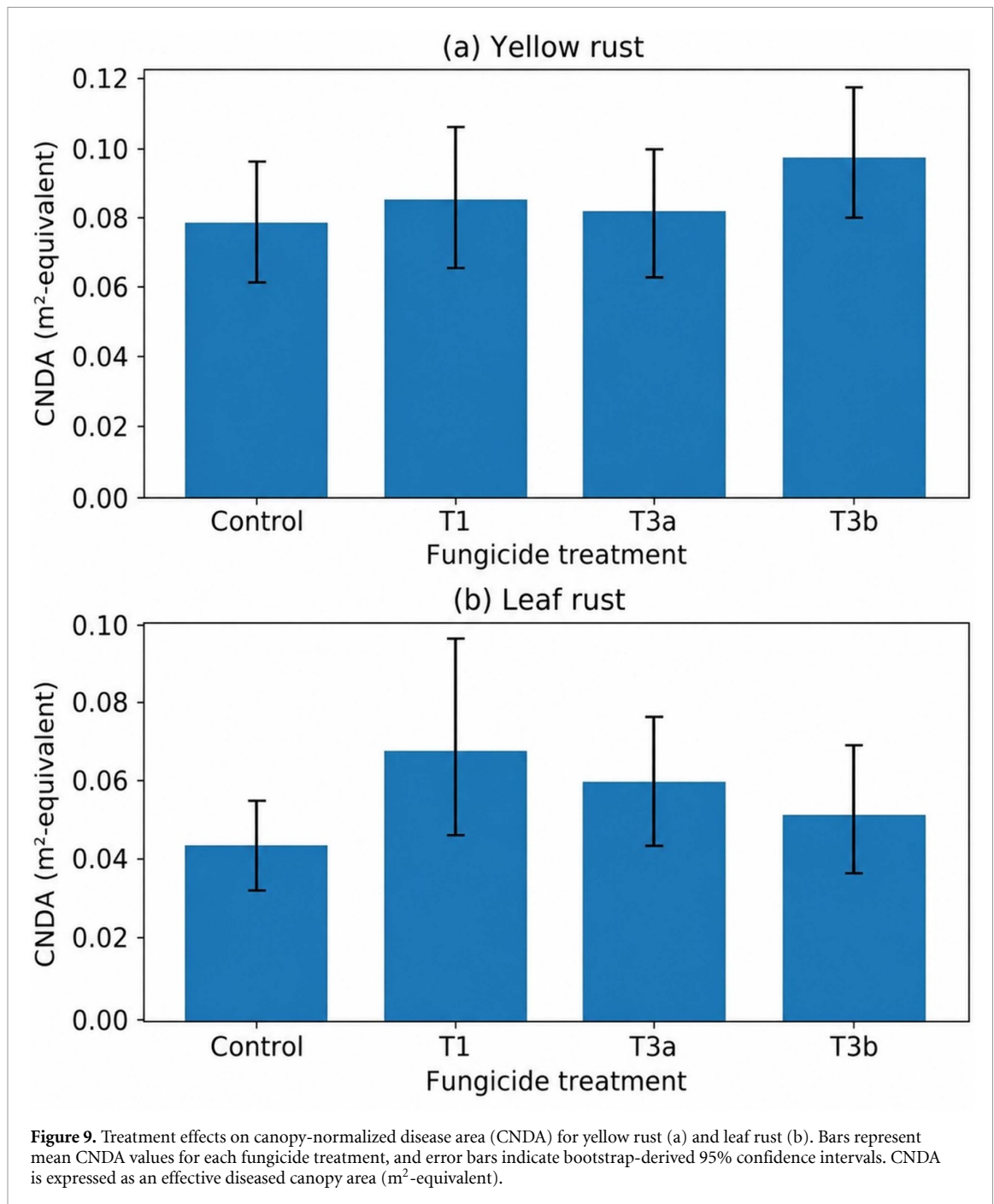
Although CNDA generated more stable treatment rankings across sites, cultivars, and canopy conditions, the substantial overlap among bootstrap-derived confidence intervals (table 3) indicates that several apparent differences among fungicide treatments should not be interpreted as statistically distinct treatment effects.

Consequently, CNDA should be viewed primarily as a framework for improving ranking consistency and effect-size interpretation across heterogeneous observational conditions rather than as a tool for definitive statistical separation among treatments.

These findings suggest that accounting for canopy-mediated observability can improve the robustness of treatment comparisons, while uncertainty estimates remain essential for avoiding over-interpretation of ranking differences.

4.6. Measurement-driven ranking shifts

Direct comparison of fungicide-treatment rankings derived from raw disease severity and the CNDA revealed disease-specific ranking shifts for both yellow rust and leaf rust (figure 10). These shifts provide empirical evidence that fungicide-induced modifications of canopy structure could substantially alter conventional severity-based treatment comparisons. Although some treatments maintained similar relative positions across methods, others exhibited marked ranking inversions after canopy normalization, indicating that canopy-mediated observability strongly influences perceived treatment performance. Importantly, the magnitude and direction of these ranking shifts differed between yellow rust and leaf rust, suggesting that the interaction between canopy structure and disease expression is pathosystem-dependent.



4.7. Statistical regimes of agreement

Agreement metrics exhibited a clear canopy-dependent evolution, with systematic changes in concordance, error magnitude, and bias across canopy cover classes (table 4). These patterns confirm the existence of distinct measurement regimes governed by canopy structure. Three dominant regimes were identified: (i) a low-canopy regime characterized by high variability and background-driven noise, (ii) a transitional regime with improving agreement, and (iii) a high-canopy regime with reduced variability but potential systematic bias. These regimes provide a conceptual framework for interpreting UAV-based disease measurements and highlight the need for canopy-aware evaluation strategies.

5. Discussion

5.1. UAV-based disease estimation as a measurement problem

Most UAV-based studies on plant disease monitoring have been framed as detection, mapping, or prediction problems, with emphasis placed on spectral feature extraction, machine learning performance,

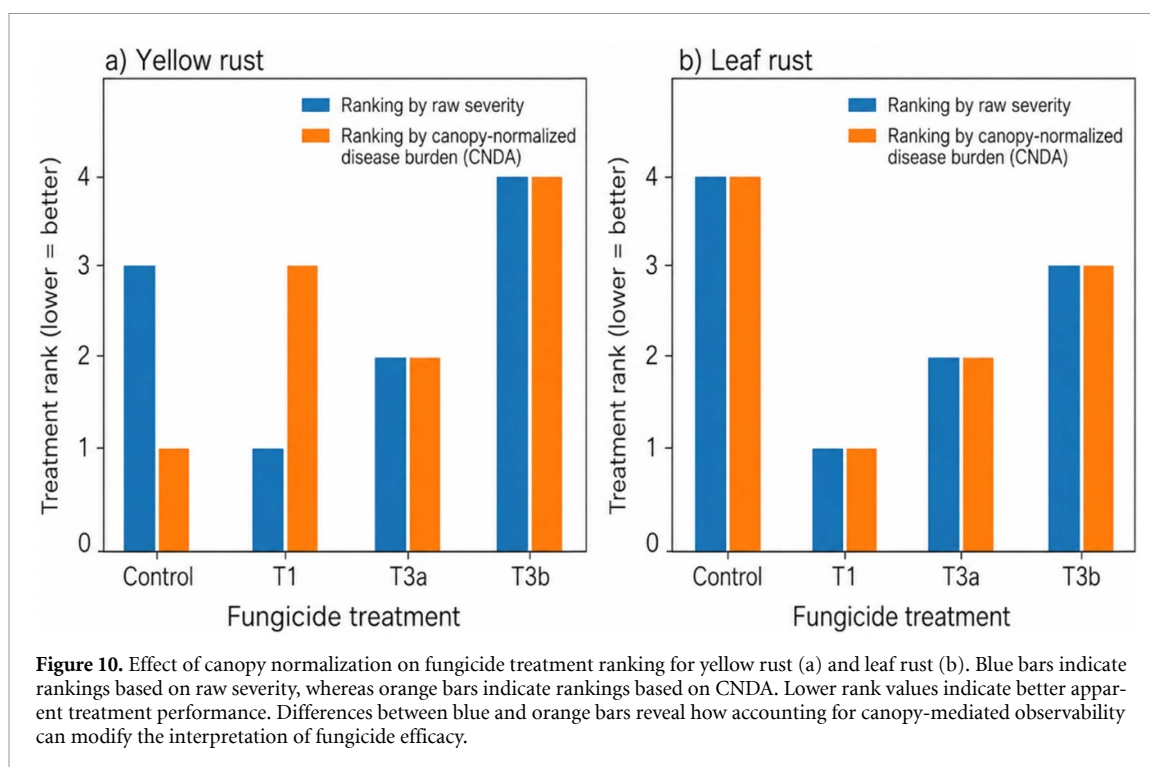


Table 4. Canopy-dependent agreement regimes between UAV- and rater-based disease assessments across yellow rust and leaf rust epidemics. Agreement, correlation, and error statistics are reported across canopy cover classes. Metrics include the Pearson's coefficient of determination (R^2), Lin's concordance correlation coefficient (CCC), mean absolute error (MAE), root mean square error (RMSE), and systematic bias (UAV–rater).

Disease	Canopy cover class	<i>n</i>	R^2 (—)	CCC (—)	MAE (%)	RMSE (%)	Bias (%)	Dominant measurement regime
Yellow rust	Low (<60%)	102	0.91	0.883	1.297	1.752	0.319	High concordance but unstable regime dominated by background noise and heterogeneous lesion visibility
	Medium (60%–80%)	126	0.87	0.823	0.854	1.225	0.112	Transitional regime with reduced bias and improving measurement stability
	High (>80%)	44	0.79	0.667	0.814	1.064	−0.185	Structurally buffered regime with stable signal but slight UAV underestimation
Leaf rust	Low (<60%)	102	0.9	0.95	2.027	3.089	0.561	Strong association but highly heteroscedastic measurement regime
	Medium (60%–80%)	126	0.73	0.681	0.84	1.287	0.378	Transitional regime with improved agreement and reduced dispersion
	High (>80%)	44	0.41	0.226	1.064	1.516	−0.801	Dense-canopy regime characterized by systematic UAV underestimation

or early symptom recognition (Mahlein 2016, Khan *et al* 2018, Gao *et al* 2025, Logavito *et al* 2025). While these approaches have substantially advanced operational capabilities, they implicitly assume a stable and unbiased relationship between observed spectral signals and disease severity. Our results suggest that the measurement regime is dynamic and influenced by management. Fungicide treatments can alter canopy structure and, in turn, affect how disease is seen by both UAV sensors and human raters. If this is ignored, models may appear accurate but not fully reflect true disease burden. This study therefore treats UAV-based disease quantification as a measurement problem that requires explicit modelling of the observation process.

Our finding aligns with broader insights from measurement science, where it is well established that high correlation or predictive accuracy does not guarantee valid measurement if the observation operator is mischaracterized (Bland and Altman 1986, Lin 1989, McClintock *et al* 2010, Morrison 2016). By separating disease processes from what can be observed, our approach links plant epidemiology with measurement theory and moves beyond purely data-driven models. It also highlights limits of detection-based methods by treating UAV-based disease estimation as a measurement process shaped by canopy-mediated observability.

These findings further support the interpretation that UAV-derived disease estimates should be considered observational outputs influenced by canopy structural conditions rather than direct measurements of the underlying biological disease process. Consequently, disease assessment should be viewed not only as a detection problem but also as a measurement problem in which canopy-mediated observability plays a central role.

5.2. Canopy cover as a state variable of disease observability

Fractional canopy cover emerged as a dominant driver of human–UAV agreement and measurement bias, acting as a state variable that controls disease observability. Under sparse canopies, soil background reflectance, heterogeneous leaf orientation, and incomplete lesion aggregation amplify uncertainty in both visual and RGB-based assessments. As canopy closure progresses, these sources of variability diminish, and the optical signal increasingly reflects an integrated canopy response. Such a pattern is consistent with studies highlighting the role of canopy structure, background effects, and illumination geometry in shaping reflectance signals (Asner 1998, Jafarbiglu and Pourreza 2023, Liu *et al* 2024, Raman *et al* 2026). Instead of separating canopy and disease effects, this study suggests that they are closely linked and together shape disease observability. In many UAV-based disease monitoring studies, canopy structure is typically regarded as a source of observational variability to be minimized. In contrast, our findings suggest that canopy cover may instead serve as an observation-state variable that influences how underlying disease burden is translated into measurable disease signals.

It should be noted that the unequal distribution of observations among canopy-cover classes may have influenced the precision and generalizability of regime-specific agreement estimates. In particular, the high-canopy class ($>80\%$ UAV_Fcover) contained substantially fewer observations ($n = 44$) than the low- and medium-canopy classes ($n = 102$ and 126 , respectively; table 2). This limitation is especially relevant for leaf rust, where the concordance correlation coefficient declined markedly under high-canopy conditions. Consequently, results associated with the dense-canopy regime should be interpreted with appropriate caution and regarded as preliminary rather than conclusive. Despite this limitation, the overall pattern remained consistent with the proposed canopy-dependent measurement framework, suggesting that canopy structure influences both UAV–rater agreement and disease observability. Further studies incorporating larger and more balanced datasets across canopy-cover classes will be important to confirm the robustness and stability of these measurement regimes.

The strong canopy dependence observed in the present study is consistent with previous reports showing that canopy architecture influences symptom visibility, illumination heterogeneity, leaf overlap, and sensor performance. These structural effects contribute to non-stationary measurement behaviour and partly explain why disease assessment accuracy may vary across epidemic stages despite similar disease severities.

5.3. Agreement-based evaluation of UAV disease estimates

A persistent limitation in the UAV imagery-based disease literature is the reliance on correlation-based validation. High Pearson or Spearman correlations are often interpreted as evidence of measurement reliability, despite long-standing warnings that correlation does not assess agreement (Bland and Altman 1986, Edwards *et al* 2021). Our results provide a clear empirical illustration of this issue: strong association between UAV- and rater-based estimates can coexist with substantial bias and heteroscedastic error. By combining Bland–Altman analysis with Lin’s CCC and stratifying these metrics by canopy cover, we showed that measurement quality was regime-dependent rather than uniform. Agreement-based diagnostics provide a useful complement to correlation metrics when interpreting UAV-derived disease estimates. This aligns UAV-based disease assessment with best practices in environmental measurement science, where agreement—not correlation—is the standard criterion for evaluating measurement validity (Schober *et al* 2018, Edwards *et al* 2021, Correia *et al* 2026).

The structured error landscapes further indicate that canopy cover and disease severity cannot be treated as independent sources of variation. Instead, both variables interact dynamically to determine symptom observability and measurement reliability, supporting the concept of canopy-dependent observational regimes.

5.4. Treatment-induced confounding and its consequences

The treatment ranking shifts provide empirical evidence that fungicide efficacy assessments depend on the measurement framework (figure 9). Fungicide strategies that keep leaves green for longer and delay senescence can change canopy structure, how disease is seen. This phenomenon represents a form of treatment-induced confounding, whereby management practices influence both the biological process of interest and the observation process itself. Similar issues have been reported in ecological monitoring, where management affects detectability rather than abundance alone (Tyre *et al* 2003, MacKenzie *et al* 2017). However, they have rarely been demonstrated quantitatively in plant disease remote sensing. Our study provides one of the first quantitative demonstrations of this confounding mechanism in UAV-based plant disease monitoring. Severity-based comparisons may therefore yield inconsistent or contradictory estimates of fungicide efficacy across sites, cultivars, or epidemic stages.

Fungicide treatments therefore affected simultaneously both the biological process (disease progression) and the observation process (disease visibility). This dual influence creates a treatment-induced observational confounding effect that may bias conventional severity-based treatment comparisons if canopy structure is not explicitly considered.

5.5. The originality of canopy-normalized disease burden

The CNDA metric represents a shift from severity-centric to burden-centric disease quantification. Severity reflects relative symptom expression but not the canopy extent or structure through which disease is expressed. By combining severity, fractional canopy cover, and plot area, the CNDA estimates an effective diseased canopy area with clearer physical and biological meaning. Importantly, CNDA should not be interpreted as a direct estimate of pathogen biomass, infection intensity, or total biological disease burden. Rather, it quantifies disease expression within the observable canopy domain available to both visual raters and UAV sensors.

Consequently, low CNDA values may arise either from genuinely low disease expression or from a reduced observable canopy domain resulting from senescence, canopy degradation, or severe disease-induced tissue loss. The metric therefore captures the interaction between disease severity and canopy structure rather than the absolute biological burden of infection.

Importantly, CNDA should not be interpreted as a new empirical index, but as a conceptual framework for harmonizing disease measurements across heterogeneous canopy conditions. Accounting for canopy-mediated observability can alter agronomic inference, particularly when evaluating fungicide efficacy. The CNDA integrates disease severity and canopy structure and was selected because it retains a clear physical interpretation as effective diseased canopy area. Spectral-index-only approaches do not capture this combined effect.

Although CNDA generated more stable treatment rankings across sites, cultivars, and canopy conditions, the overlap among bootstrap confidence intervals indicates that several apparent differences among fungicide treatments should not be interpreted as statistically distinct treatment effects. Consequently, CNDA primarily supports ranking consistency and effect-size interpretation rather than definitive treatment separation.

The principal contribution of CNDA is therefore not to provide a biologically superior measure of disease burden, nor to replace conventional disease severity assessments, but rather to reduce canopy-mediated observational distortions that may influence treatment comparisons. By accounting for structural differences among canopies, CNDA improves canopy-aware observational comparability across heterogeneous observation conditions.

The value of CNDA lies principally in reducing observational confounding rather than increasing statistical power. Within this framework, CNDA should be viewed as an observational harmonization tool that improves canopy-aware comparability rather than a replacement for biological disease metrics.

5.6. Implications and future research pathways

Operationally, UAV-based disease monitoring systems should incorporate canopy-aware uncertainty information rather than relying exclusively on deterministic disease severity maps. Canopy cover may serve as an indicator of measurement reliability and could be integrated into decision-support frameworks to flag potentially uncertain disease estimates, refine intervention thresholds, or trigger complementary field observations. These findings support the development of disease monitoring systems that are both data-driven and measurement-aware by explicitly accounting for canopy-related uncertainty during disease quantification and management decisions.

An important implication of this study is that disease quantification should not be viewed solely as a detection problem but also as a measurement problem. In UAV-based disease monitoring, observed

disease severity reflects not only the underlying epidemiological process but also the structural conditions governing symptom observability. Canopy architecture, leaf overlap, senescence dynamics, background reflectance, illumination heterogeneity, and symptom visibility collectively influence how disease symptoms are expressed and detected by both human observers and remote sensing systems. Consequently, disease estimates may contain a structural component linked to observability in addition to the biological signal itself.

Within this framework, the CNDA provides a means of integrating disease severity with canopy structure and spatial context. Importantly, CNDA should be interpreted as a canopy-aware observational metric rather than as a direct measure of pathogen biomass, infection intensity, or absolute biological disease burden. The metric does not quantify pathogen population size, lesion density, infection efficiency, sporulation potential, or epidemic intensity *per se*. Instead, it quantifies disease expression within the observable canopy domain available to both visual raters and UAV sensors, thereby reducing structural biases associated with differences in canopy development and canopy closure.

Accordingly, the principal value of CNDA lies in improving canopy-aware comparability among treatments exposed to contrasting structural conditions rather than providing a biologically superior measure of disease. The metric was specifically designed to harmonize disease observations across heterogeneous canopy states and to reduce canopy-mediated observational distortions that may influence conventional severity-based comparisons.

This distinction is particularly important when evaluating fungicide efficacy. Successful fungicide applications frequently modify canopy structure simultaneously with disease progression through delayed senescence, maintenance of green leaf area, and changes in canopy architecture (Robert *et al* 2008, Vidal *et al* 2017). Consequently, part of the apparent treatment effect observed through conventional disease severity assessments may originate from changes in symptom observability rather than from disease suppression alone. Accounting for canopy-mediated observability may therefore improve the interpretation of treatment performance and strengthen the robustness of UAV-based disease assessment frameworks.

The treatment-ranking analyses further illustrate this point. CNDA generated treatment rankings that were generally more consistent across sites, cultivars, and canopy conditions than rankings derived solely from disease severity. However, these ranking differences should be interpreted cautiously. Although CNDA frequently improved ranking consistency and clarified effect-size direction, overlapping bootstrap confidence intervals indicate that several treatment differences cannot be considered statistically distinct. Consequently, CNDA should primarily be viewed as a tool for improving ranking stability and effect-size interpretation rather than as evidence of definitive statistical separation among fungicide treatments.

More generally, the value of CNDA lies principally in reducing observational confounding associated with canopy structure rather than increasing statistical power. Improvements in treatment comparability do not necessarily imply stronger treatment effects but instead reflect a reduction in canopy-mediated measurement distortions that may influence disease quantification.

The present study was conducted under a specific set of experimental conditions involving one growing season, four field sites, two wheat cultivars, RGB imagery, and two rust pathosystems. Although the observed relationships were highly consistent across sites and diseases, caution is warranted when extrapolating these findings beyond the conditions investigated here. Broader validation across multiple growing seasons, additional cultivars, wider disease severity ranges, different environmental conditions, and alternative disease pathosystems will be necessary before the proposed framework can be considered universally transferable.

Future research should therefore investigate the applicability of canopy-aware measurement frameworks across a broader range of crops, diseases, and sensing technologies. Particular attention should be given to evaluating whether similar canopy-mediated observability effects occur in pathosystems characterized by different symptom architectures, disease distributions, and epidemic dynamics. Additional studies should also assess the robustness of CNDA using multispectral, hyperspectral, thermal, and emerging sensor platforms capable of capturing complementary information on plant physiological status and disease development.

Furthermore, because canopy classes were not equally represented in the present dataset, agreement statistics obtained under near-complete canopy closure (>80% canopy cover) should be interpreted with appropriate caution. Nevertheless, the observed decline in UAV–visual agreement under dense canopy conditions was consistently detected across both yellow rust and leaf rust epidemics, suggesting that the identified canopy-dependent measurement regime reflects a genuine biological and observational phenomenon rather than a sampling artefact. Future datasets containing larger numbers of observations under fully closed-canopy conditions would further strengthen the robustness of these conclusions.

Ultimately, these results suggest that future advances in UAV-based disease monitoring may depend not only on improvements in sensor performance or machine-learning algorithms but also on a better

understanding of the biological and structural factors that govern disease observability. Integrating measurement theory, canopy structure, and epidemiological processes may therefore provide a more robust foundation for next-generation disease monitoring systems in precision agriculture.

5.7. Limitations and implications

While canopy normalization improved the comparability of disease assessments across treatments and sites, the CNDA framework remains influenced by severe canopy deterioration during advanced epidemic stages. In heavily affected untreated plots, accelerated senescence and canopy collapse reduced the observable canopy domain available for symptom expression and detection. As a result, low CNDA values under sparse-canopy conditions may not necessarily indicate a low underlying disease burden but may instead reflect reduced symptom observability arising from structural canopy loss.

This limitation underscores an important distinction between latent disease burden and observed disease signals. UAV-derived disease estimates are influenced not only by pathogen activity but also by the structural and optical conditions that determine symptom visibility. Consequently, the CNDA formulation integrates both epidemiological and canopy-structural components of disease expression. Although this feature represents a conceptual strength of the framework, it also implies that CNDA should be interpreted as an observationally conditioned indicator of disease burden rather than a direct physical measure of infected tissue.

Importantly, CNDA should not be interpreted as a biologically superior measure of disease burden, pathogen biomass, infection intensity, or epidemic severity. Rather, its primary value lies in providing a canopy-aware observational framework that reduces structural distortions affecting disease quantification and improves comparability among treatments exposed to contrasting canopy conditions.

Similarly, although CNDA frequently produced more stable treatment rankings than raw severity assessments, overlapping bootstrap confidence intervals indicate that some treatment differences should not be interpreted as statistically distinct. Consequently, CNDA is particularly valuable for improving ranking consistency and effect-size interpretation rather than for establishing definitive statistical separation among treatments.

The value of CNDA therefore lies principally in reducing observational confounding associated with canopy-mediated symptom visibility rather than in increasing statistical power or replacing conventional epidemiological measurements. In this sense, CNDA should be viewed as an observational harmonization tool that facilitates canopy-aware interpretation of disease expression while remaining complementary to biological disease metrics.

Another limitation concerns the scope of the present dataset. The framework was evaluated during a single growing season across four experimental sites, two wheat cultivars, two rust pathosystems, and using RGB imagery exclusively. Although these conditions generated substantial variability in canopy structure, disease intensity, and fungicide response, broader validation across additional seasons, cultivars, pathosystems, environmental conditions, and sensing technologies remains necessary before the framework can be considered fully transferable.

Furthermore, canopy cover classes were not equally represented in the dataset. Consequently, agreement statistics associated with dense-canopy conditions should be interpreted with appropriate caution. Nevertheless, the overall trends remained remarkably consistent across both yellow rust and leaf rust, suggesting that the identified canopy-mediated observability regimes reflect genuine measurement processes rather than sampling artefacts.

Future research could explore alternative normalization approaches, including regression-based correction methods, hierarchical latent-state frameworks, or Bayesian observability models that explicitly disentangle biological disease burden from canopy-mediated visibility effects. Future studies should also evaluate the applicability of canopy-aware disease metrics under contrasting climatic conditions, additional cereal diseases, diverse agronomic management systems, and alternative sensing technologies including multispectral, hyperspectral, and thermal platforms. Such investigations will help determine whether canopy-mediated observability constitutes a general property of plant disease measurement across crop–pathogen–sensor combinations.

Incorporating observability processes into epidemiological analyses may ultimately enhance the robustness, interpretability, and transferability of UAV-based disease monitoring across heterogeneous field environments. More broadly, these findings suggest that plant disease quantification should be considered not only as a detection challenge but also as a measurement-science problem, in which canopy structure, symptom visibility, and sensor characteristics jointly influence the reliability of disease observations.

6. Conclusion

This study demonstrates that canopy structure is not a secondary factor but a fundamental determinant of disease observability in UAV-based disease monitoring. Our results suggest that fungicide treatments may influence both disease dynamics and the measurement process itself, indicating that conventional severity metrics may not always fully capture treatment responses when canopy structure differs substantially among treatments.

The CNDA framework helps mitigate this source of observational bias by explicitly accounting for canopy-mediated filtering, thereby improving the consistency and comparability of disease observations across heterogeneous canopy conditions. Importantly, CNDA should not be interpreted as a direct estimate of pathogen biomass, infection intensity, or total biological disease burden. Rather, it represents a canopy-aware observational metric designed to characterize disease expression within the observable canopy domain available to both visual raters and UAV sensors.

The results further indicate that the principal contribution of CNDA lies in reducing canopy-mediated observational confounding rather than increasing statistical power or providing definitive separation among fungicide treatments. By improving canopy-aware comparability, the framework facilitates more robust interpretation of treatment responses while preserving appropriate caution regarding statistical uncertainty.

More broadly, these findings highlight the value of considering plant disease quantification as a measurement-science challenge rather than solely a detection problem. Disease observations are influenced not only by underlying epidemic processes but also by structural factors governing symptom visibility and detectability. Incorporating these mechanisms into UAV-based monitoring frameworks may improve the robustness, interpretability, and transferability of remote sensing approaches for agricultural decision-making.

Nevertheless, the present study was conducted during a single growing season using RGB imagery across four sites, two cultivars, and two wheat rust pathosystems. Additional validation across multiple years, crops, diseases, environmental conditions, and sensing technologies will therefore be required before canopy-aware disease quantification frameworks can be considered broadly transferable across agricultural systems.

Overall, the proposed framework should be viewed as a proof-of-concept demonstrating the importance of canopy-mediated observability in disease assessment and highlighting the potential of canopy-aware observational harmonization for improving the interpretation of UAV-derived disease measurements.

Data availability statement

All data that support the findings of this study are included within the article (and any supplementary files).

Supplementary data available at <https://doi.org/10.1088/2515-7620/ae85e2/data1>.

Author contributions

Moussa El Jarroudi  0000-0002-9169-8824

Conceptualization (equal), Data curation (equal), Formal analysis (equal), Funding acquisition (equal), Investigation (equal), Methodology (equal), Project administration (equal), Resources (equal), Software (equal), Validation (equal), Visualization (equal), Writing – original draft (equal), Writing – review & editing (equal)

Louis Kouadio  0000-0001-9669-7807

Conceptualization (equal), Data curation (equal), Formal analysis (equal), Methodology (equal), Software (equal), Supervision (equal), Validation (equal), Writing – review & editing (equal)

Jonathan Peereman  0000-0002-5128-3536

Writing – review & editing (equal)

Marco Beyer  0000-0002-9415-4718

Funding acquisition (equal), Project administration (equal), Writing – review & editing (equal)

References

- Anam I, Arafat N, Hafiz M S, Jim J R, Kabir M M and Mridha M F 2024 A systematic review of UAV and AI integration for targeted disease detection, weed management, and pest control in precision agriculture *Smart Agric. Technol.* **9** 100647
- Asner G P 1998 Biophysical and biochemical sources of variability in canopy reflectance *Remote Sens. Environ.* **64** 234–53
- Baccar R, Fournier C, Dornbusch T, Andrieu B, Gouache D and Robert C 2011 Modelling the effect of wheat canopy architecture as affected by sowing density on Septoria tritici epidemics using a coupled epidemic-virtual plant model *Ann. Bot.* **108** 1179–94
- Bland J M and Altman D G 1986 Statistical methods for assessing agreement between two methods of clinical measurement *Lancet* **327** 307–10
- Bock C H, Chiang K-S and Del Ponte E M 2022 Plant disease severity estimated visually: a century of research, best practices, and opportunities for improving methods and practices to maximize accuracy *Trop. Plant Pathol.* **47** 25–42
- Bock C H, El Jarroudi M, Kouadio L A, Mackels C, Chiang K-S and Delfosse P 2014 Disease severity estimates—effects of rater accuracy and assessment methods for comparing treatments *Plant Dis.* **99** 1104–12
- Bundessortenamt 2018 Beschreibende Sortenliste. Getreide, Mais, Öl- und Faserpflanzen, Leguminosen, Rüben und Zwischenfrüchte; Bundessortenamt: hannover, Germany. ISSN 2190–6130 (available at: www.bundessortenamt.de/bsa/media/Files/BSL/bsl_getreide_2018.pdf) (Accessed 28 May 2026)
- Chin R, Catal C and Kassahun A 2023 Plant disease detection using drones in precision agriculture *Precis. Agric.* **24** 1663–82
- Correia H E *et al* 2026 Best practices for moving from correlation to causation in ecological research *Nat. Commun.* **17** 1981
- Deery D M and Jones H G 2021 Field phenomics: will it enable crop improvement? *Plant Phenomics* **2021** 9871989
- Edwards C, Allen H and Chamunyonga C 2021 Correlation does not imply agreement: a cautionary tale for researchers and reviewers *Sonography* **8** 185–90
- Fones H N, Bebbler D P, Chaloner T M, Kay W T, Steinberg G and Gurr S J 2020 Threats to global food security from emerging fungal and oomycete crop pathogens *Nat. Food* **1** 332–42
- Gao J, Gujarati K, Hegde M, Arra P, Gupta S and Buch N 2025 Integration of UAV and remote sensing data for early diagnosis and severity mapping of diseases in maize crop through deep learning and reinforcement learning *Remote Sens.* **17** 3427
- Gilligan C A 2007 Sustainable agriculture and plant diseases: an epidemiological perspective *Phil. Trans. R. Soc. B* **363** 741–59
- Heidarian Dehkordi R, El Jarroudi M, Kouadio L, Meersmans J and Beyer M 2020 Monitoring wheat leaf rust and stripe rust in winter wheat using high-resolution UAV-based red-green-blue imagery *Remote Sens.* **12** 3696
- Jafarbiglu H and Pourreza A 2023 Impact of sun-view geometry on canopy spectral reflectance variability *ISPRS J. Photogramm. Remote Sens.* **196** 270–86
- Jeger M J, Pautasso M, Holdenrieder O and Shaw M W 2007 Modelling disease spread and control in networks: implications for plant sciences *New Phytol.* **174** 279–97
- Khan I, Liu H, Cheng T, Tian Y, Cao Q, Zhu Y, Cao W and Yao X 2018 Detection of wheat powdery mildew based on hyperspectral reflectance through SPA and PLS-LDA *Int. J. Precis. Agric. Aviat.* **1** 43–48
- Kouadio L, El Jarroudi M, Belabess Z, Laasli S-E, Roni M Z K, Amine I D I, Mokhtari N, Mokrini F, Junk J and Lahlali R 2023 A review on UAV-based applications for plant disease detection and monitoring *Remote Sens.* **15** 4273
- Lin L I-K 1989 A concordance correlation coefficient to evaluate reproducibility *Biometrics* **45** 255–68
- Liu W, Möttus M, Gastellu-etchegorry J-P, Fang H and Atherton J 2024 Seasonal and vertical variation in canopy structure and leaf spectral properties determine the canopy reflectance of a rice field *Agric. For. Meteorol.* **355** 110132
- Logavito G, Horanont T, Thapa A, Intarat K, Wuttiwong K-O and Sun X 2025 Field-scale detection of bacterial leaf blight in rice based on UAV multispectral imaging and deep learning frameworks *PLoS One* **20** e0314535
- MacKenzie D I, Nichols J D, Royle J A, Pollock K H, Bailey L and Hines J E 2017 *Occupancy Estimation and Modeling: Inferring Patterns and Dynamics of Species Occurrence* (Academic) (<https://doi.org/10.1016/C2012-0-01164-7>)
- Mahlein A-K 2016 Plant disease detection by imaging sensors—parallels and specific demands for precision agriculture and plant phenotyping *Plant Dis.* **100** 241–51
- Mahlein A-K, Arnal Barbedo J G, Chiang K-S, Del Ponte E M and Bock C H 2024 From detection to protection: the role of optical sensors, robots, and artificial intelligence in modern plant disease management *Phytopathology* **114** 1733–41
- Manoj H M, Shanthi D L, Lakshmi B N, Archana K J, Venkata Naga Jyothi E and Archana K 2025 AI-driven drone technology and computer vision for early detection of crop disease in large agricultural areas *Sci. Rep.* **16** 2479
- McClintock B T, Nichols J D, Bailey L L, MacKenzie D I, Kendall W L and Franklin A B 2010 Seeking a second opinion: uncertainty in disease ecology *Ecol. Lett.* **13** 659–74
- Milne A, Paveley N, Audsley E and Livermore P 2003 A wheat canopy model for use in disease management decision support systems *Ann. Appl. Biol.* **143** 265–74
- Morrison L W 2016 Observer error in vegetation surveys: a review *J. Plant Ecol.* **9** 367–79
- Peterson R F, Campbell A B and Hannah A E 1948 A diagrammatic scale for estimating rust intensity on leaves and stems of cereals *Can. J. Res.* **26c** 496–500
- R Core Team 2024 *R: A Language and Environment for Statistical Computing* (R Foundation for Statistical Computing) (available at: www.R-project.org/)
- Raman R, Neely H L, Rajan N, Siegfried J, Ibrahim A M H, Adams C B and Hardin R G 2026 Soil background effects on UAS and proximal remote sensing-derived vegetation indices *Agron. J.* **118** e70281
- Robert C, Fournier C, Andrieu B and Ney B 2008 Coupling a 3D virtual wheat (*Triticum aestivum*) plant model with a Septoria tritici epidemic model (Septo3D): a new approach to investigate plant-pathogen interactions linked to canopy architecture *Funct. Plant Biol.* **35** 997–1013
- Savary S, Wilcoquet L, Pethybridge S J, Esker P, McRoberts N and Nelson A 2019 The global burden of pathogens and pests on major food crops *Nat. Ecol. Evol.* **3** 430–9
- Schober P, Boer C and Schwarte L A 2018 Correlation coefficients: appropriate use and interpretation *Anesth. Analg.* **126** 1763–8
- Shafay M, Hassan T, Owais M, Hussain I, Khawaja S G, Seneviratne L and Werghi N 2025 Recent advances in plant disease detection: challenges and opportunities *Plant Methods* **21** 140
- Singh R P *et al* 2025 Challenges to wheat disease resistance and current global strategies *Annu. Rev. Phytopathol.* **63** 201–24
- Tyre A J, Tenhumberg B, Field S A, Niejalke D, Parris K and Possingham H P 2003 Improving precision and reducing bias in biological surveys: estimating false-negative error rates *Ecol. Appl.* **13** 1790–801
- Vidal T, Boixel A-L, Durand B, de Vallavieille-pope C, Huber L and Saint-Jean S 2017 Reduction of fungal disease spread in cultivar mixtures: impact of canopy architecture on rain-splash dispersal and on crop microclimate *Agric. For. Meteorol.* **246** 154–61