Luis Kremer
Faculty for Environmental and Natural Resources
Chair of Sensor-based Geoinformatics
Tennenbacher Str. 4, 79106 Freiburg, Germany

Feb 13th 2026

egusphere-2025-5223: AngleCam V2: Predicting leaf inclination angles across taxa from daytime and nighttime photos

Dear Handling Associate Editor Mirco Migliavacca, Dear Referees,

We thank both referees for their careful review of our manuscript and appreciate the constructive comments that have helped improve the quality of our work. Each concern and technical correction has been considered, and the text has been revised accordingly. Below is a point-by-point response to all comments:

Text with a vertical grey line marks the original referee's comment.

Text with a vertical dark blue line is our response. Line numbers indicated in these paragraphs refer to the original manuscript.

Text with a vertical light blue line indicates our suggested changes to the original manuscript. Line numbers here refer to the revised manuscript (without tracked changes), which will be made available.

The main improvements of the manuscript include:

- Separation of model performance between daytime (RGB) and nighttime (NIR) conditions within the multitemporal terrestrial laser scanning (TLS) validation experiments.
- Enlarged pseudo-NIR examples in the Appendix, including a qualitative evaluation with actual night-vision imagery.
- Refinement of the Discussion to address the capabilities and constraints of the pseudo-NIR implementation, which enables the model to process nighttime images. Additionally, replacing general claims of "reliability" with specific performance metrics.
- Inclusion of a bin-wise distribution error analysis assessing the model's capacity to capture the complete leaf inclination angle distribution (LIAD) shape, moving beyond average leaf angles.

We hope that the revised manuscript meets the requirements of the referees and the journal, and we look forward to your response.

Sincerely,
Luis Kremer
(on behalf of the co-authors)

Response to comments by Referee #1

Leaf angle distribution is an important canopy structural variable, but its temporal and spatial dynamics are much less well understood compared to LAI. The primary reason is that we don't have a reliable method to measure it at a large spatial scale yet. This work developed AngleCam V2, which uses a deep learning model to estimate leaf inclination angle distributions from RGB and night-vision NIR imagery. By expanding training data across species, canopy structures, and lighting conditions, as well as integrating synthetic near-infrared augmentation, the developed model shows improved generalization and temporal robustness, including nighttime compatibility, which is promising for large-scale applications. The current version is well written and clearly presents the advantage of the new version of AngleCam. My only concern is the accuracy of the training datasets, in terms of visual interpretation, 20 leaf samples, representativeness of the entire canopy.

> We thank the referee for the helpful comments and constructive review. Please find our detailed, point-by-point responses to each concern below.

Vertical grey line: Original referee's comment.

Vertical dark blue line: Our response (line numbers refer to the original manuscript).

Vertical light blue line: Revisions to the manuscript text (line numbers refer to the revised manuscript).

Major concerns

Referee comment 1.1: Section 2.1.2, L182-185: It is unclear that how the authors sampled the 20 leaves? Please provide more details, especially the vertical distribution of leaf samples. Why did the authors use the 20 leaf samples here? How are the estimates sensitive to the leaf sample number?

> **Response:** Selecting 20 leaves per image balances capturing leaf angle variability within each image and building a diverse training dataset for robust generalization across species, canopy structures, scenes, and lighting conditions. While annotating more leaves per image would reduce sampling uncertainty, it would also increase labeling time and limit the number of images in the training set. We therefore prioritize including more images and conditions, while maintaining consistent sampling within each image.
>
> The 20-leaf sampling protocol follows AngleCam V1 (Kattenborn et al., 2022), where the underlying labels supported successful validation of the model against terrestrial

laser scanning (R^2 = 0.74). To ensure representative sampling, we apply an evenly spaced point-grid overlay and annotate the leaf at each selected grid point. This approach enforces spatial balance across the field of view and reduces selection bias toward prominent or easily interpretable leaves. With a horizontal camera orientation, the field of view includes leaves at multiple canopy heights, allowing the sampled leaves to capture a vertical gradient within the visible canopy.

We add further details on the creation of the reference data set in the Methods:

Section 2.1.2, L183-L190: We obtained a LIAD for each reference image by sampling the average leaf surface inclinations of 20 leaves. This sample size was chosen to obtain a representative intra-image leaf angle variability as well as a large image dataset across species and scene conditions. We followed the approach from AngleCam V1, where this sampling strategy was validated by terrestrial laser scanning (R^2 = 0.74) (Kattenborn et al., 2022). For each image, we placed an evenly spaced grid of points and selected the leaf closest to each point for annotation. This method ensures spatial balance across the field of view, reduces bias toward prominent leaves, and samples leaves throughout the vertical gradient of the visible canopy seen from the horizontal camera angle.

Referee comment 1.2: Section 2.1.2: It is unclear to me how the authors determined the leaf angle for each sample? If the estimates were made through visual inspection, how do the authors account for potential uncertainties and the possible differences among observers?

Response: Leaf inclination angles are determined by visual interpretation of horizontal RGB images (see L179-L181), as this approach enables efficient annotation of a large number of samples. We agree that visual interpretation can introduce observer-related uncertainty and the risk of systematic bias.

To reduce sensitivity to uncertainty at the label level, individual angle measurements are not used directly. Instead, beta distributions are fitted to the 20 annotated angles per image, and distribution augmentation is applied during training (±20% of the standard deviation), which increases the model's robustness to plausible measurement variation and sampling uncertainty (see also Kattenborn et al., 2022). Additionally, the model is trained on thousands of annotated images, so random interpretation errors are expected to average out during optimization rather than systematically propagate into the learned signal.

The capability of machine learning and deep learning to overcome noisy training data is reported in a range of studies (see e.g., review by Kattenborn et al., 2024), but still, careful evaluation is necessary. To verify that AngleCam V2 does not simply reproduce potential annotation artefacts and biases, the model is evaluated not only against visually interpreted reference data but also against independent leaf angle measurements derived from terrestrial laser scanning. This external validation provides confidence that the model's performance reflects actual canopy leaf-angle patterns rather than systematic effects of the manual labeling procedure.

For the same Section (2.1.2), we make the following change to address the concerns:

Section 2.1.2, L191-209: Although this approach enables efficient annotation, it may introduce observer-related uncertainty. To reduce sensitivity to label uncertainty, individual discrete angle measurements were not used directly as training targets. Instead, the 20 leaf samples were converted into probability distributions across the full 0°–90° leaf angle range. The two-parameter beta distribution was employed, as it is well-suited to model continuous gradients of leaf inclination angle distribution shapes, ranging from strongly horizontal (planophile) to uniform, spherical, or strongly vertical (erectophile) forms (Goel & Strebel, 1984). Beta distribution fitting was performed in Python using the scipy.stats module (v1.13.0), with parameters (α, β) estimated by maximum likelihood.

Rather than directly predicting the (α, β) parameters of a beta distribution, the model was trained to predict the entire probability density function (PDF) across the 0°–90° range at 2° intervals, since the underlying LIAD may differ from an idealized beta distribution in some cases (Kattenborn et al., 2022). To further mitigate the impact of potential annotation bias and enhance generalization, the fitted (α, β) values of the reference LIADs were augmented, generating 50 synthetic variants for each image by sampling (α, β) within ±20% of the standard deviation of the maximum likelihood estimates. This data augmentation increases the model's robustness to plausible measurement variations. Additionally, by training on over 3500 images, random interpretation errors are expected to be mitigated during optimization rather than systematically propagating into the learned signal (Rolnick et al., 2017). The robustness of this labeling approach was verified by evaluating the model against independent leaf angle measurements derived from terrestrial laser scanning (see Section 2.3).

Referee comment 1.3: Whether do the horizontal photographs can represent the entire canopy? If the photos are taken from different azimuths, whether the estimated LIADs will be different?

Response: We indeed did not elaborate on this in the original manuscript. A single horizontal photograph samples only the foliage visible within the camera's field of view, typically at a distance of approximately 1–2 meters from the camera, and thus represents a local portion of the canopy rather than the entire canopy structure. This design ensures that AngleCam reports the LIAD of the visible leaf surfaces from the camera perspective, adhering to a "what you see is what you get" principle.

Concerning the azimuthal effects, LIAD estimates can vary when photographs are taken from different compass directions. However, these differences may reflect genuine canopy asymmetries or directional patterns in leaf positioning, which AngleCam is designed to capture. Therefore, camera placement and field design should explicitly consider representativeness and potential azimuthal variability. As shown in Kattenborn et al. (2022), a spatial replication, including images acquired from different azimuths and different heights, can be employed to develop a more comprehensive characterization of the variability of leaf angle dynamics within the canopy.

We add the following paragraph in the Discussion to elaborate on these topics:

Section 4.4, L572-L581: When deploying AngleCam for ecological monitoring, it is essential to consider the spatial representativeness of the imagery. A single horizontal photograph captures only the foliage within the camera's field of view, typically representing a local portion of the canopy at a distance of approximately 1 to 2 meters. As a result, AngleCam reports the LIAD of the visible leaf surfaces from that specific perspective. Because leaf orientation may display directional patterns or canopy asymmetries, such as those caused by phototropism or prevailing winds, LIAD estimates can vary depending on the azimuth from which photographs are taken. To achieve a representative characterization of the entire canopy structure, camera placement should account for these spatial variations. We recommend spatial replication, including image acquisition from multiple azimuths and different heights, to capture a comprehensive distribution of leaf angles.

Referee comment 1.4: Section 1.3: Please clearly present the reasonability of the pseudo-NIR image generation, as well as its uncertainties and associated impacts on the developed model.

Response: The integration of grayscale and pseudo-NIR images into the training pipeline addresses the modality gap encountered when processing both RGB and true NIR imagery (see L209-L210). We cannot apply a model trained on 3-channel RGB on monochromatic NIR imagery. However, creating ample amounts of reference data for the NIR imagery would have been logistically very expensive. Given the abundance of existing RGB labels and the intensive effort required to generate a comparable amount of NIR annotations, we chose to leverage our existing dataset by transforming RGB images into synthetic proxies. These transformations are designed to replicate key characteristics of real NIR images, such as their monochromatic appearance, high vegetation reflectance, and distance-dependent illumination falloff.

Specifically, we acknowledge that the pseudo-NIR pipeline is an approximation that retains daytime ambient shadows and may exhibit depth-based overexposure relative to true NIR backscatter. However, independent and successful validation against terrestrial laser scanning indicates that the model generalizes to true night-vision environments, achieving a nighttime RMSE of 3.1° to 3.6° for the tested species. Although this approach enables continuous 24-hour monitoring, we expand Section 4.3 in the Discussion and add Section A.2 in the Appendix to provide an overview of its capabilities and limitations.

Referee comment 1.5: Section 2.1.4: Please provide more details on how to determine the parameters in the deep learning models.

Response: We update Section 2.1.4 (L251–256) to detail the tested hyperparameter ranges, the final configuration, and the hardware and time used for training.

Section 2.1.4, L262-272: The model was trained on an NVIDIA RTX A6000 GPU (48 GB RAM), with the final training run of 50 epochs completed in 24 minutes. Hyperparameters were determined through testing of different configurations, including batch sizes (8–64), optimizers (Adam, AdamW, and SGD), and learning rates (10^{-5} to 10^{-1}), alongside different regression head architectures with varying dropout rates (0–0.5) and weight decay settings (10^{-5} to 10^{-1}).

The final configuration used the AdamW optimizer with an initial learning rate of 1×10^{-4}, weight decay of 0.01, and gradient clipping (max norm = 3.0, L2 norm) to prevent overfitting (Loshchilov & Hutter, 2017). Learning rate scheduling was managed with ReduceLROnPlateau (reduction factor 0.5, patience 5 epochs, minimum learning rate 1×10^{-6}). Training was performed with a batch size of 32 and Huber loss to reduce sensitivity to outliers (Sun et al., 2020). The final model was selected based on the lowest validation loss obtained through the 50 epochs.

Referee comment 1.6: Section 3.1: I am curious about whether the model performance depends on the leaf angle distribution of vegetation itself. For example, the model may show better performance on spherical vegetation that other types.

Response: Yes, the model performance depends on the distribution of leaf inclination angles. Our error analysis shows that prediction errors are not significantly linked to taxonomy (phylogenetic signal across 100 genera: Pagel's $\lambda \approx 0$; Moran's I < 0.1), so differences in performance are not primarily due to relatedness between species. Instead, the model tends to overestimate low average leaf inclination angles and underestimate high average leaf inclination angles. This matches regression-to-the-mean effects from the training data, which has mostly intermediate angles and fewer extreme cases below 20° or above 60° (see L408-411). As a result, the model works best for intermediate LIAD types (like spherical or plagiophile) and is less accurate for extreme planophile or erectophile types (see also Fig. 3).

We add this consideration in the Discussion:

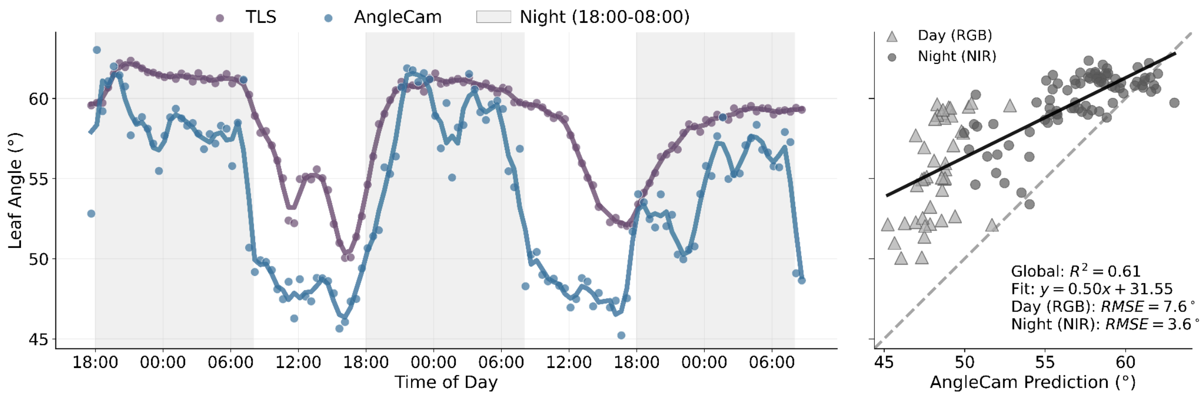
Section 4.1, L452-L454: Consequently, AngleCam V2 achieves its highest performance for intermediate LIAD types, such as spherical or plagiophile, while demonstrating reduced accuracy for extreme planophile or erectophile canopy structures. [...]

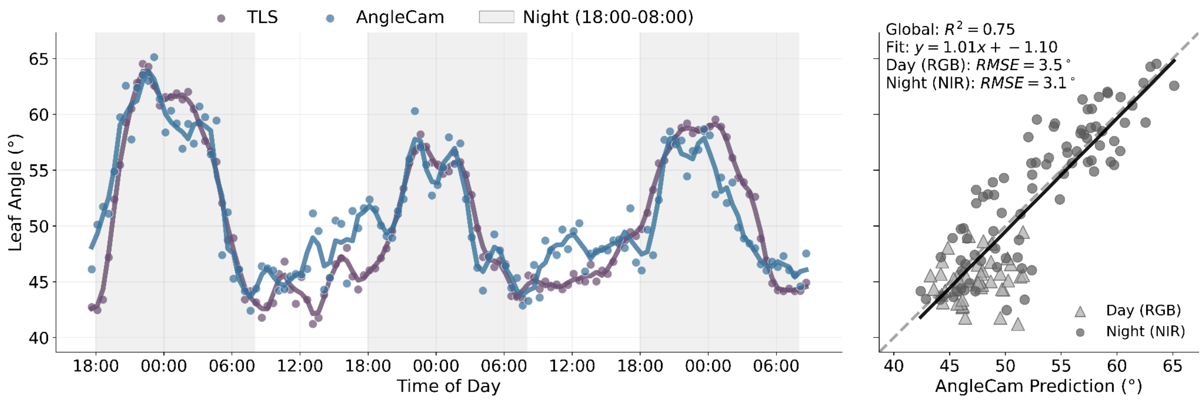
Referee comment 1.7: Section 3.3: Does the model performance vary between day and night?

Response: Important point that was not addressed before. We revise the corresponding figure to explicitly separate performance metrics for daytime (RGB) and nighttime (NIR) conditions, with the Root Mean Square Error (RMSE) now reported for each modality. We also expand Section 3.3 to include these results:

Section 3.3, L393-395: [...] The average leaf angles derived from the two methods yielded a high global correspondence (R^2 = 0.61). Disaggregated by modality, the Root Mean Square Error (RMSE) was 7.6° for daytime (RGB) and 3.6° for nighttime (NIR) imagery.

Section 3.3, L396-400: The results for Maranta leuconeura showed stronger agreement. [...] The global correspondence was R^2 = 0.75. The model achieved close alignment across both modalities, with an RMSE of 3.5° for daytime (RGB) and 3.1° for nighttime (NIR) imagery.





Referee comment 1.8: Is it possible to estimate leaf angle distribution using photos taken from different angles (e.g., PhenoCam) based on the current modeling framework?

> **Response**: Accurate estimation of leaf inclination angles using the current AngleCam framework requires images to be captured with a horizontal camera orientation. If the camera is tilted, as is often the case with PhenoCams, the apparent leaf angles in the image shift relative to gravity, resulting in biased absolute angle estimates. Therefore, the current model cannot be directly applied to oblique imagery without introducing additional uncertainty.
>
> We clarify this in the Discussion:
>
> *Section 4.4, L565-L568: [...] However, accurate LIAD estimation with AngleCam depends on a horizontal camera orientation. Successful integration into these networks would require either identifying existing cameras with appropriate horizontal geometries or installing dedicated sensors that conform to the required acquisition geometry. [...]*

Minor concerns

Referee comment 1.9: Data Availability Statement: Please provide the brief introduction on the shared data in the data availability section. Please also introduce the file naming and folder structure in the repo.

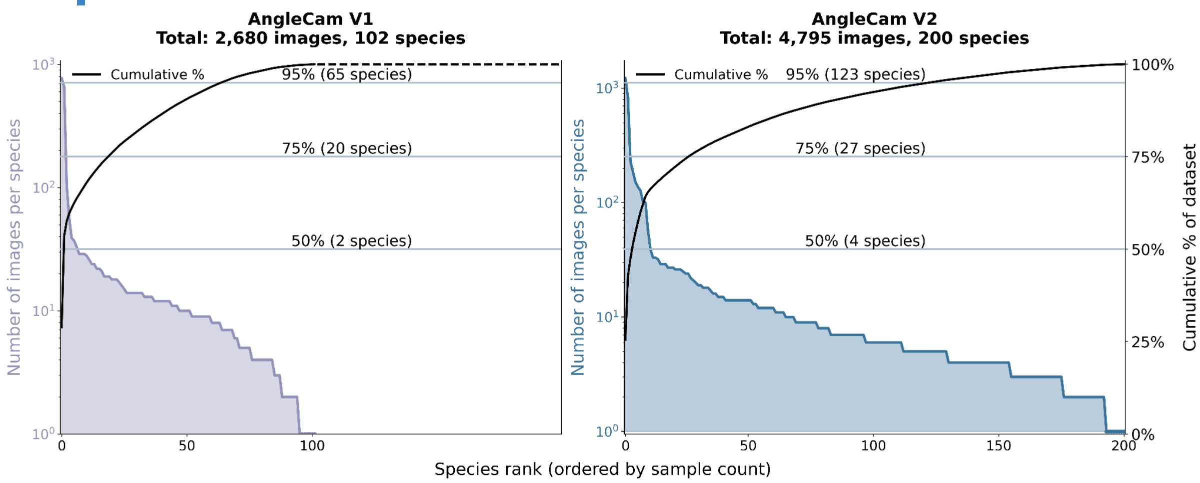
> **Response**: Thank you for the suggestion. We expand the data availability statement to briefly describe the shared data and its contents. The actual project repository is now referenced, with an explanation of the dataset folder structure.

We make the following changes to the data availability statement:

All data are available on Zenodo at https://doi.org/10.5281/zenodo.17086253, organized into four directories: training/validation images with annotations (01_Training_Validation_Data/), multitemporal test imagery with TLS validation (02_Test_Data/), model weights and predictions (03_Model_Outputs/), and supplementary materials (04_Supplementary_Material/). The source code is available at https://github.com/Luis-Kr/AngleCamV2. The pretrained model is available at https://doi.org/10.5281/zenodo.17101166.

Referee comment 1.10: Figure 1: Please enlarge the font size. For the right panel, the y-axis tick labels are missing.

Response: Thank you for pointing this out. We increase the font size in Figure 1 and add the missing y-axis tick labels in the right panel.



Referee comment 1.11: Does the model performance depend on the distance at which the photo is taken?

Response: Camera-to-canopy distance can affect model performance by altering the visible foliage and detail in the image. The training data includes images taken at varying distances within the typical acquisition range (generally less than 2 meters), so AngleCam V2 is expected to be robust within this range, although a sensitivity analysis for distance was not conducted. In practice, the relevant distance range also depends on leaf size, as very small leaves become difficult to resolve at greater distances.

We clarify this in the Discussion:

> *Section 4.4, L569–L571: [...] Similarly, the distance to the canopy influences the effective spatial resolution. While the model training covered typical detection ranges (< 2 m), images must be taken close enough to clearly resolve individual leaves.*

Referee comment 1.12: Line 486: canhelp -> can help

> **Response**: Thank you for pointing out this typo. We correct “canhelp” to “can help” (Line 486).

Referee comment 1.13: Can AngleCam capture the impacts of wind on leaf angle variation?

> **Response**: Wind can induce rapid leaf movements that may affect instantaneous AngleCam estimates. With high-frequency image acquisition (e.g., 1–2 min), this variability can be captured in the time series and may even be used as an environmental proxy (see Kattenborn et al., 2022). If the objective is to emphasize slower trends, such as diurnal or drought-driven changes, aggregating predictions to coarser intervals (e.g., 3 h, 6 h, 12 h) helps to smooth wind effects.

> We thank the referee for taking the time to review the manuscript and the constructive feedback!

Response to comments by Referee #2

> This is a useful and timely preprint that makes a clear methodological advance on AngleCam V1 by greatly expanding taxonomic and site diversity, analyzing phylogenetic structure in errors, and validating temporal dynamics against TLS. The methods are generally transparent, and open code/data availability is excellent. But in the present form, it would merit revisions focused on sharpening and modestly tempering the headline claims, and clarifying evaluation design and limitations, especially for the pseudo-NIR.

> > We thank the referee for the helpful comments and constructive review. Please find our detailed, point-by-point responses to each concern below.

> Vertical grey line: Original referee’s comment.

> Vertical dark blue line: Our response (line numbers refer to the original manuscript).

> Vertical light blue line: Revisions to the manuscript text (line numbers refer to the revised manuscript).

General comments

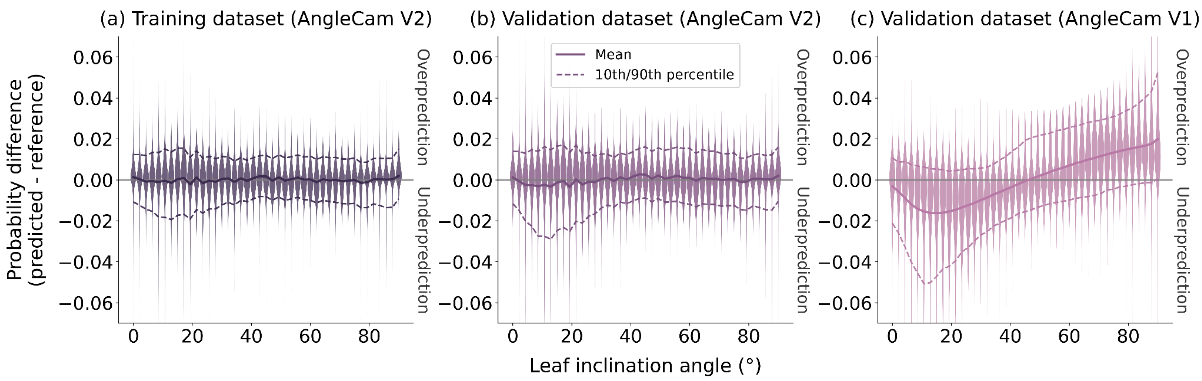
> **Referee comment 2.1:** The Introduction devotes substantial space to LIADs and motivates the work around recovering full leaf angle distributions. However, in the Results the primary “headline” performance reporting is largely based on average leaf angle, with limited quantitative assessment of the predicted distribution shape. This creates a mismatch between what is promised early and what is actually presented. I recommend either: (1) adding distribution-level evaluation metrics (e.g., alpha, beta metrics or simply the skewness you briefly discussed) and reporting them prominently alongside mean-angle metrics; or (2) reframing the Introduction to emphasize mean angle as the primary validated output and positioning LIAD recovery as a secondary/ongoing capability. Without one of these changes, readers may reasonably expect stronger LIAD-level validation than is currently provided.

> > **Response:** Thank you for this helpful observation. We agree that the manuscript currently emphasizes average leaf angle in the Results, while the Introduction motivates the recovery of full LIADs. To address this mismatch, we add a distribution-level evaluation in the manuscript that shows uncertainties across the full LIAD range (0-90°). Specifically, we include a figure in Section 3.1 showing the per-bin bias (reference minus prediction) across all 43 angle bins for both the training and validation datasets. We considered using skewness as a primary metric, but it can be overly sensitive to small errors in the tails. Therefore, we focus on bin-wise

distribution errors, which provide a more direct and interpretable assessment of LIAD shape.

To clarify, AngleCam V2 does not directly predict the parameters of a beta distribution. Instead, the model outputs a 43-bin discretized LIAD (0-90° in 2° steps), represented as a flexible one-dimensional probability vector (see L191-192). During training, these predictions are supervised using reference distributions derived from beta fits, but the model output is not constrained to be beta-shaped. Therefore, evaluation based solely on fitted beta parameters (such as α and β) would introduce an additional processing step (re-fitting a beta distribution to the predicted vector) and would not necessarily capture the full distributional differences.

Section 3.1: Fig. 4: Bin-wise prediction errors for leaf inclination angle distributions are shown for: (a) the training dataset (AngleCam V2, n = 3591), (b) the validation dataset (AngleCam V2, n = 899), and (c) the validation dataset (AngleCam V1, n = 899). Each plot displays the error distribution (predicted minus reference probability) for every 2° angle bin. Solid lines show the mean error, while dashed lines represent the 10th and 90th percentiles. Positive values indicate overprediction and negative values underprediction.



Section 3.1, L363-372: In addition to average angle performance, the model's ability to capture the entire distribution shape was evaluated by analyzing bin-wise prediction errors (predicted minus reference probability) across 43 leaf angle bins (Fig. 4). For the training set, the mean error remained centered around zero for all angle bins, with a stable percentile band and only a minor negative dip in the 10th percentile between 10° and 25° (Fig. 4a). In the validation set, AngleCam V2 similarly maintained a mean error near zero, with a slight positive deviation between 40° and 60°, and a more pronounced negative dip in the 10th percentile at lower angles (10°–25°) (Fig. 4b). By contrast, AngleCam V1 exhibited a pronounced angle-dependent bias on the same validation data, with underprediction at low angles (10°–20°), increasing overprediction at higher angles, and a wide distribution spread toward extreme orientations (Fig. 4c).

Section 4.1, L443-451: [...] We assume that this bias, also known as regression to the mean bias (Barnett et al., 2005), emerges from the training data distribution, which contains abundant samples between 35°-50° but relatively few extreme cases *below 20° or above 60°. The bin-wise error analysis (Fig. 4) further clarifies this bias, where AngleCam V2 maintains a mean error near zero across the distribution range but shows an increased spread in the 10th percentile at lower angles (10°–25°) in the validation set. This pattern reflects greater uncertainty when the model encounters these less frequently represented orientations. Nevertheless, the stability of the mean error across all bins represents a major advancement over AngleCam V1, which exhibited a pronounced, systematic angle-dependent bias and increasing variance toward extreme orientations.*

Referee comment 2.2: Nighttime/NIR capability is not validated broadly enough to support strong claims. The paper introduces pseudo-NIR and mixed-modality training, but "validation was performed exclusively on original RGB images" and NIR generalization is inferred largely from limited indoor experiments. The conclusions claim reliable tracking "across day and night imagery," which should be softened unless supported by broader true-NIR validation across more scenes/species/cameras. The manuscript should also more explicitly characterize what is and isn't captured by the pseudo-NIR pipeline and provide ablations showing its benefit on real NIR test imagery.

Response: We agree that the current manuscript wording could be interpreted as claiming broad, scene- and camera-independent performance on true nighttime/NIR imagery. We implement the following changes to address the concerns:

(1) We revise the Abstract, Introduction, and Conclusions to emphasize the pseudo-NIR approach as a way to address the labeling constraints for real NIR images (see L209-213). We replace general claims of "reliability" with specific performance metrics from the TLS experiments.

(2) We expand Section 4.3 in the Discussion and add Appendix Section A.2 to explicitly detail what the pipeline captures (e.g., illumination falloff, blue-sky darkening) and what it misses (e.g., true NIR-LED shadow behavior, backscatter intensity nuances).

(3) We report performance metrics (RMSE) separately for daytime RGB and true-nighttime NIR imagery in the Results (Section 3.3) and the corresponding figure. These results provide a direct quantitative assessment of the model's performance on real NIR test imagery.

Abstract, L48-51: [...] To address the lack of labeled NIR data, a training strategy was developed that uses pseudo-NIR imagery derived from daytime RGB images. The model is based on a vision transformer architecture with mixed-modality training incorporating RGB, grayscale, and pseudo-NIR images.

Abstract, L54-57: [...] Testing against leaf angle dynamics from multitemporal terrestrial laser scanning confirmed the model's ability to track diurnal leaf

> *movements (R^2 = 0.61–0.75) with a nighttime RMSE of 3.1° – 3.6° (compared to daytime RMSEs of 3.5° and 7.6°, respectively). [...]*

> *Section 1, L110-113: [...] (3) we developed a training strategy that combines grayscale and pseudo-NIR transformations to leverage existing RGB labels and minimize manual effort for NIR annotation. This approach enables the model to process both RGB and NIR night-vision imagery, supporting continuous 24-hour monitoring of leaf movements.*

> *Section 4.3, L487-505: [...] This capability emerged from our dual-modality training strategy, combining grayscale conversion with pseudo-NIR generation as a proxy for actual night-vision data. This workaround was required by the high labor intensity of manual leaf annotation. While our RGB dataset is extensive, replicating this sample size with manually labeled NIR imagery was not feasible within the scope of the current study.*
>
> *As shown by the TLS comparison, this strategy proved effective for the tested species. For Calathea ornata (Global R^2 = 0.61), the model achieved an RMSE of 7.6° for daytime (RGB) and 3.6° for nighttime (NIR) imagery. For Maranta leuconeura (Global R^2 = 0.75), performance was even more consistent, with an RMSE of 3.5° for RGB and 3.1° for NIR imagery. Although these results indicate that the model can generalize to real night-vision environments, the pseudo-NIR pipeline remains an approximation with inherent physical limitations. Specifically, pseudo-NIR images retain daytime ambient shadows and may exhibit depth-based overexposure or less plausible illumination falloff compared to true NIR backscatter (a detailed characterization of these discrepancies and visual comparisons to real NIR imagery are provided in Appendix Section A.2). As a result, the synthetic labels do not serve as a perfect substitute for true NIR annotations. Nevertheless, this method effectively leverages the existing RGB dataset to enable 24-hour monitoring, avoiding the constraints of extensive manual NIR labeling. Future iterations incorporating large-scale, independently labeled NIR datasets will be essential to further improve night-vision accuracy and to address sensor-specific discrepancies.*

> *Conclusions, L611-614: [...] Expanding the training dataset to include diverse species and canopy structures, combined with grayscale and pseudo-NIR transformations that leverage extensive existing RGB labels, has improved generalization and temporal robustness, including compatibility with nighttime imagery.*

Specific comments

Referee comment 2.3: L66-121: The ecological motivation is compelling but dense, with many overlapping citations about leaf angles and energy balance. Consider trimming slightly and re-allocating space to more explicit “What’s new in AngleCam V2?”.

> **Response:** Thank you for this suggestion. We streamline the ecological motivation in the Introduction (L71–L91) by condensing the background on leaf orientation and its importance. We then highlight more explicitly the novelties of AngleCam V2, specifically: (1) significantly expanded taxonomic diversity in training and validation, (2) the implementation of a vision transformer (DINOv2) for improved feature extraction, and (3) a training strategy combining grayscale and pseudo-NIR transformations to leverage existing RGB labels.

> *Section 1, L70-L86: Even within a single plant, leaf angles vary to optimize resource use. Upright upper leaves facilitate light transmission, while horizontal lower leaves maximize light absorption (Niinemets, 2010). This variation, which has been studied since Darwin (Darwin & Darwin, 1883), regulates light uptake, competition, microclimates, and canopy productivity (Hikosaka & Hirose, 1997; Mullen et al., 2006; Niinemets, 2010).*
>
> *The vertical orientation of a leaf is defined by its surface angle, which ranges from 0° (entirely horizontal) to 90° (entirely vertical). The variability of these angles within a canopy is referred to as the leaf inclination angle distribution (LIAD). LIADs are not static, as plants adjust leaf angles through active and passive processes, including diurnal movements, phenological changes, and water-stress-induced drooping. These responses are influenced by light, temperature, competition, and plant water status (Apelt et al., 2017; Geldhof et al., 2021; Niinemets, 2010; Puglielli et al., 2017). Rapid changes in leaf angles can indicate stress, such as heat or photoinhibition, and regulate canopy energy balances (Ehleringer et al., 1987; Kattenborn et al., 2022; Leuzinger & Körner, 2007; Sastry et al., 2018; Van Zanten et al., 2010; Werner et al., 2001). LIAD variability also significantly influences canopy reflectance, making it critical for global vegetation monitoring and interpreting satellite-based signals (Braghiere et al., 2021; Dechant et al., 2020; Jablonski et al., 2025; Kattenborn et al., 2024).*

> *Section 1, L106–L113: Here, we present AngleCam V2, which addresses these limitations through three major advancements: (1) we expanded the training and validation dataset to over 4500 images across 200 species, ensuring robust generalization across diverse growth forms and biomes. (2) we integrated a self-supervised vision transformer (DINOv2) to enhance feature extraction and robustness to varied scene conditions. (3) we developed a training strategy that combines grayscale and pseudo-NIR transformations to leverage existing RGB*

> *labels and minimize manual effort for NIR annotation. This approach enables the model to process both RGB and NIR night-vision imagery, supporting continuous 24-hour monitoring of leaf movements.*

> **Referee comment 2.4:** L106: “The plausibility and potential of AngleCam were underlined, given that predicted leaf inclination angle distributions over multiple months had a tight relationship to environmental conditions”. It is not convincing to show the plausibility by comparing with environmental conditions. There are cases where leaf angle was decoupled with those environmental conditions. The TLS method serves as a more convincing validation. It is also helpful to explain why you only did a single TLS scan, given the possible occlusion you mentioned in line 452. Multiple scans around the plant can overcome this issue, providing better reference data.

> **Response:** We acknowledge that the Introduction implies that validation relies substantially on correlations with environmental conditions. We revise this sentence to present the long-term time series from AngleCam V1 as an application example, demonstrating plausible dynamics and sensitivity to environmental drivers, rather than positioning it as the main validation.
>
> Regarding the TLS reference data, we agree that using multiple scan positions can reduce occlusion and lead to more complete reference LIADs. In our multitemporal experiment, scans were acquired every 30 minutes over multiple day-night cycles, and repositioning the TLS during this continuous acquisition was not feasible. To ensure the highest possible reference quality under these constraints, we used a high-precision instrument (Riegl VZ-400i) and optimized its position for maximum coverage. We include a statement in the Discussion that future experiments could benefit from multi-position TLS, such as multiple fixed scanners or sequential multi-view scans in shorter campaigns, to reduce occlusion effects and improve reference quality.

> *Section 1, L100-L103: [...] The applicability and responsiveness of AngleCam were demonstrated through long-term time series in which predicted leaf inclination angle distributions exhibited plausible dynamics in response to environmental drivers, including radiation, temperature, and soil humidity at multiple sites (Kattenborn et al., 2022, Kattenborn et al., 2024). [...]*

> *Section 4.3, L509-512: [...] However, the fixed TLS position, which was necessary for continuous 30-minute sampling over several days, likely introduced some occlusion-related bias. Future studies could address this by employing multiple fixed scanners or, for shorter intensive campaigns, by conducting sequential multi-view scans.*

> **Referee comment 2.5:** L172–L181: The labeling method is described as visual estimation of whole leaf surface inclination. Please clarify how occluded/curled leaves are handled, whether the 20 leaves are sampled across canopy strata, and whether annotators used any geometric cues (e.g., midrib line) consistently.

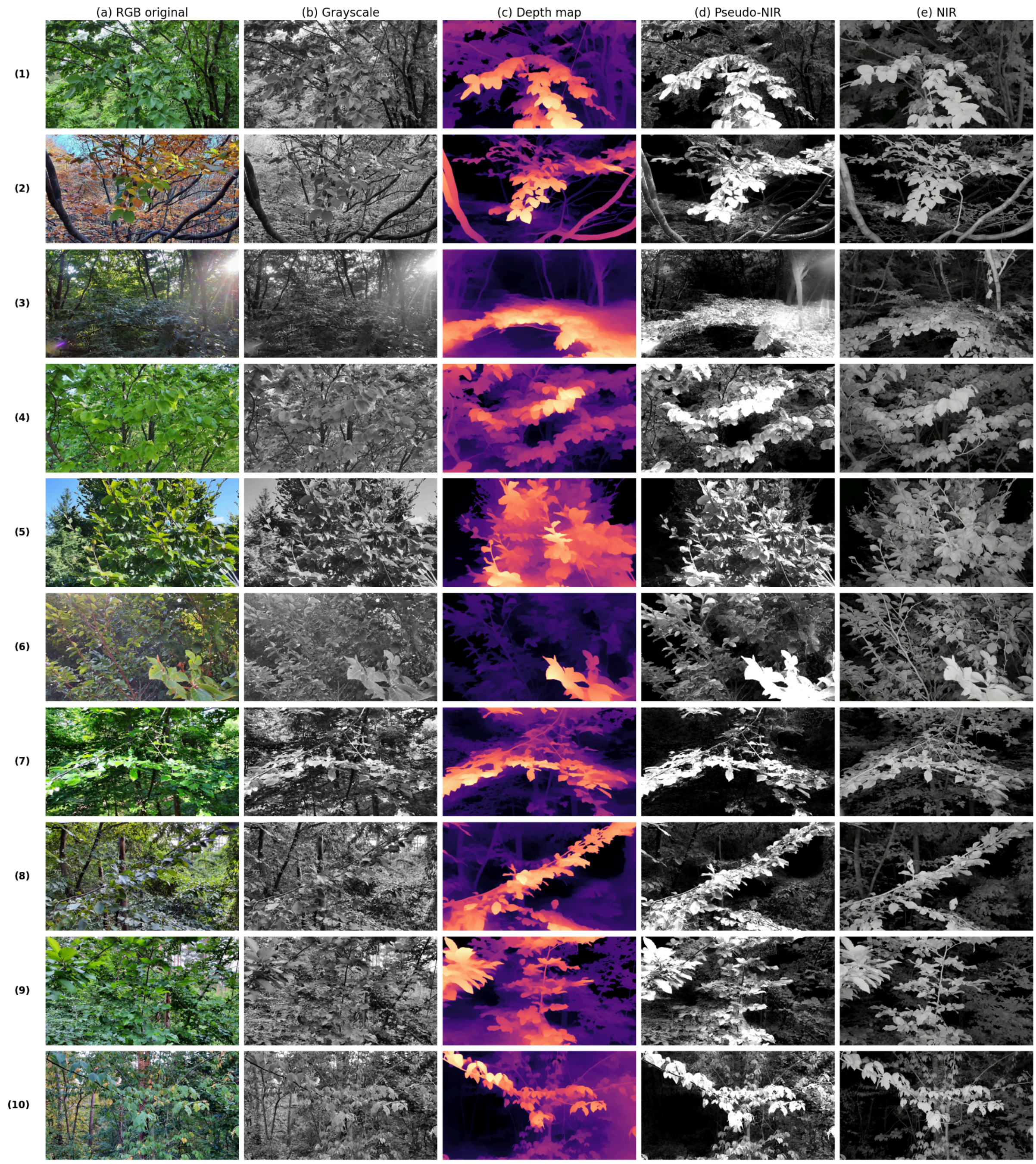
> **Response:** Thank you for addressing this shortcoming. We state this more clearly in the Methods. Leaf inclinations are estimated from horizontal RGB images by visual interpretation only, without added geometric cues. Only clearly visible leaves are annotated, and occluded leaves are not considered. Curled or non-planar leaves are included if visible, and annotators estimate the overall average leaf inclination for those samples. For each image, we select 20 leaves using an evenly spaced grid to ensure sampling across the field of view, which often covers foliage at several heights. The section is changed as follows:

> *Section 2.1.2: L177-L182: [...] Here, we overcome these limitations by estimating the average leaf inclination of whole leaf surfaces through visual interpretation of their apparent angle from horizontal in RGB images. This process relied solely on visual estimation, without additional geometric cues. Only clearly visible leaves were annotated, while occluded leaves were excluded. For curled or non-planar leaves, annotators estimated the average inclination of the entire leaf surface. This approach was previously applied successfully for AngleCam V1 (Kattenborn et al., 2022).*

> *Section 2.1.2: L187-190: [...] For each image, we placed an evenly spaced grid of points and selected the leaf closest to each point for annotation. This method ensures spatial balance across the field of view, reduces bias toward prominent leaves, and samples leaves throughout the vertical gradient of the visible canopy seen from the horizontal camera angle.*

> **Referee comment 2.6:** L199–229: Again, the synthetic NIR pipeline is clearly described, but its adequacy as a proxy for real night-vision images is not quantitatively assessed beyond qualitative Fig. 2. Consider adding a supplemental figure comparing pseudo-NIR and real NIR across several species/backgrounds showing similarities and known discrepancies.

> **Response:** We add a characterization in Appendix Section A.2 with Supplementary Figure A1. This section provides a side-by-side visual comparison of the pseudo-NIR pipeline against true night-vision imagery across ten different forest scenes. We detail specific discrepancies, such as the retention of daytime shadows in pseudo-NIR images and differences in illumination falloff compared to true NIR backscatter.



> **Referee comment 2.7:** L244–249: If possible, report separate performance metrics for daytime RGB vs pseudo-NIR in the TLS experiments serving as the main evidence for NIR robustness.

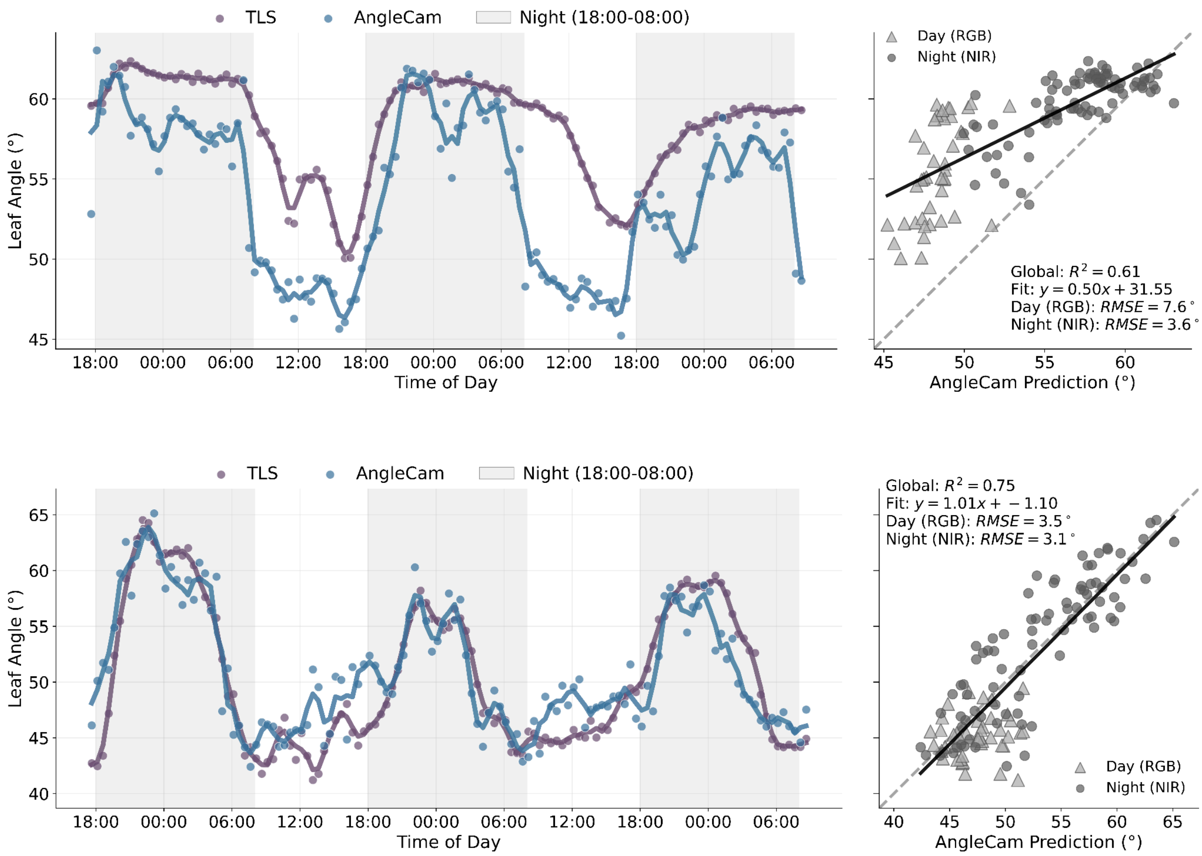
> **Response:** We update Section 3.3 and the corresponding figure to report modality-specific performance. For Maranta leuconeura, the nighttime (NIR) RMSE is 3.1° compared to a daytime (RGB) RMSE of 3.5°. For Calathea ornata, the nighttime RMSE is 3.6° compared to 7.6° for daytime. These results confirm that the model is

applicable to true night-vision imagery with error margins comparable to daytime RGB predictions.

We revise the corresponding figure to explicitly separate performance metrics for daytime (RGB) and nighttime (NIR) conditions, with the Root Mean Square Error (RMSE) now reported for each modality. We also expand Section 3.3 to include these results:

Section 3.3, L393-395: [...] The average leaf angles derived from the two methods yielded a high global correspondence (R^2 = 0.61). Disaggregated by modality, the Root Mean Square Error (RMSE) was 7.6° for daytime (RGB) and 3.6° for nighttime (NIR) imagery.

Section 3.3, L396-400: The results for Maranta leuconeura showed stronger agreement. [...] The global correspondence was R^2 = 0.75. The model achieved close alignment across both modalities, with an RMSE of 3.5° for daytime (RGB) and 3.1° for nighttime (NIR) imagery.



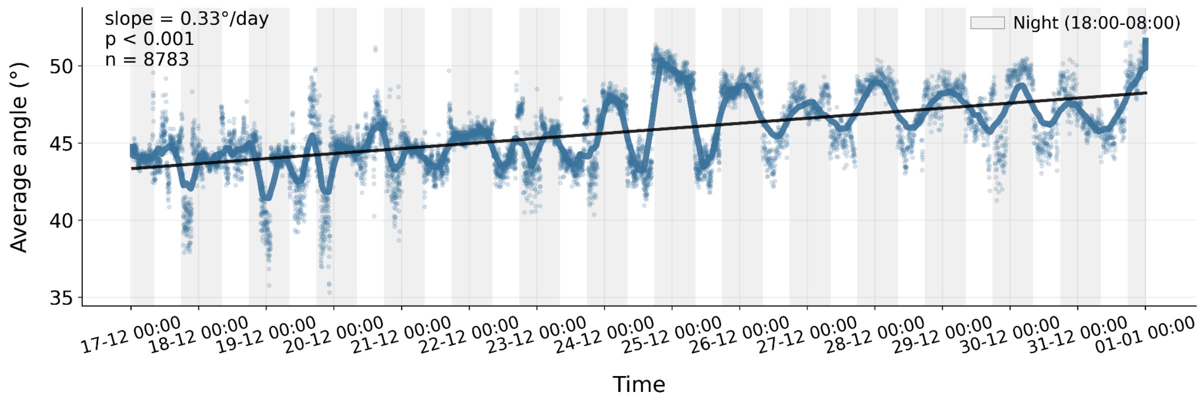
Referee comment 2.8: L380–392: You nicely contrast V1 vs V2 on their respective validation sets and on the new dataset. Please clarify whether V1 was re-trained on the expanded dataset or simply applied as originally trained. Current wording could be interpreted either way.

Response: AngleCam V1 was not retrained on the expanded dataset and was used as originally trained, then applied to the same validation dataset for comparison. We clarify this in the manuscript:

Section 4.1, L417-423: [...] However, when both models were evaluated on the new, more globally representative dataset, AngleCam V2 demonstrated substantially superior generalization, with an R² of 0.62, compared to AngleCam V1's R² of 0.12. For this comparison, AngleCam V1 was applied as originally trained without any retraining on the expanded dataset. The increased performance can be attributed to the fact that AngleCam V2 was trained on an expanded dataset with considerably greater species diversity and scene heterogeneity than the first version. [...]

Referee comment 2.9: L370–379: The 5° net increase over 14 days (0.33°/day) is modest relative to the ~9° RMSE reported for validation. It would be valuable to include at least a p-value for the trend line to show that the long-term change is statistically significant from day-to-day variability. Please also clarify whether any concurrent measurements of soil moisture or plant water status were taken; currently the inference that changes are drought-driven rests solely on watering history and visual interpretation.

Response: We add the p-value ($p < 0.001$) for the linear trend in both the Results and Fig. 5 to confirm that the long-term increase is statistically significant. No concurrent measurements of soil moisture or plant water status were collected. We add a statement in the Methods to clarify this.



Section 2.4, L346-349: [...] The plant was last watered on December 12 and thus remained without irrigation for 19 days by the end of monitoring. While the experiment followed this known watering history, no concurrent measurements of soil moisture or plant water status were collected.

Section 3.4, L404-406: [...] The time series revealed a consistent increase in average leaf angle from approximately 43° to 48° (slope = 0.33°/day, $p < 0.001$, $n = 8783$), becoming especially evident after December 25. [...]

> **Referee comment 2.10:** L268–283: The genus-level approach is sensible. But please indicate whether Pl@ntNet identifications were manually checked for obvious mislabels.

> **Response:** Yes, Pl@ntNet identifications were manually checked before being included. We assessed the plausibility of each identification at both the species and genus levels. While identification at the species level based on images may be subject to greater uncertainty, classification at the genus level is generally more robust. We therefore performed phylogenetic error analysis at the genus level as a conservative choice. We clarify this in the manuscript:

> *Section 2.2, L293-L297: For images where species identification was not available in the original metadata, we determined plant species using the Pl@ntNet online identification tool (PlantNet, 2025). Each identification was manually reviewed to assess plausibility at both species and genus levels. Due to increased uncertainty in species-level assignments from images, phylogenetic error analysis was performed at the genus level. [...]*

> **Referee comment 2.11:** L344–353; 418–437: You conclude that there is "no phylogenetic structure" and "no systematic phylogenetic bias". Given the relatively small sample of 100 genera, noisy MAE estimates, and possible genus-ID uncertainty, it would be more cautious to write "no statistically significant phylogenetic structure was detected in the residuals." It may be worth noting that absence of detectable structure does not preclude modest clade-specific biases that could appear with larger sample sizes.

> **Response:** We agree with the referee. We revise the Results (L345–353) and Discussion (L419–438). We now state that no statistically significant phylogenetic structure is detected in the residuals (Pagel's λ near zero; Moran's I non-significant). We also note that the absence of a detectable signal in our sample (100 genera) does not rule out modest clade-specific biases that might be revealed by larger sample sizes or higher taxonomic resolution.

> *Section 3.2, L378-383: [...] When mapped onto the phylogenetic tree, prediction errors showed no significant clustering. Pagel's λ was near zero ($\lambda = 7.41 \times 10^{-5}$), with no significant deviation from the null hypothesis of $\lambda = 0$ (no phylogenetic structure, $p = 1.0$), indicating that no statistically significant phylogenetic structure was present in the residuals. Consistently, Moran's I showed weak positive autocorrelation ($I = 0.074$, $p = 0.23$), supporting the finding that model residuals were not significantly structured by evolutionary relatedness within the sampled 100 genera.*

Section 4.2, L464-468: [...] Pagel's λ was near zero ($\lambda = 7.41 \times 10^{-5}$, $p = 1.0$), indicating that no statistically significant phylogenetic structure was detected in the prediction errors. Similarly, Moran's I showed only weak, non-significant autocorrelation ($I = 0.074$, $p = 0.23$), providing no evidence of systematic phylogenetic bias within the current dataset.

Section 4.2, L479-482: [...] Overall, while the absence of a detectable phylogenetic signal suggests that within-genus variation in leaf morphology and growth dynamics is reasonably captured by our training set, we acknowledge that the limited sample size (100 genera) may mask modest clade-specific biases that could emerge with larger sample sizes or higher taxonomic resolution.

Technical corrections

Referee comment 2.12: L144–149 Bullet list: add or remove terminal periods for consistency

Response: Thank you for pointing this out. We standardize the punctuation in the bullet list.

Referee comment 2.13: L244–245: "AngleCamV2" (without space) vs "AngleCam V2" elsewhere; choose one format and use consistently.

Response: We change "AngleCamV2" to "AngleCam V2"

Referee comment 2.14: L370–373: "From this point onward, angles consistently exceeded baseline values…". Consider specifying "pre-December 25 baseline" for clarity.

Response: We agree with this suggestion and clarify this in the text to refer to the pre-December 25 baseline.

Referee comment 2.15: Fig. 5 caption: "An interactive visualization can be assessed at Anonymous GitHub.". Typo: change "assessed" to "accessed"

Response: Thank you for pointing this out. We change "assessed" to "accessed".

Referee comment 2.16: L488: "deployment capability canhelp to 488 address this knowledge gap." Typo: add a space between "canhelp"

Response: We fix the typo in the text.

> We thank the referee for taking the time to review the manuscript and the constructive feedback!