

Drone + RGB Pilot Study Report

August 26, 2025, Hayato Nakanishi

Introduction

Previous research established a reference-based measurement method under controlled laboratory conditions. This report presents the findings from a foundational field experiment designed to bridge the gap between the lab and practical application. The primary goals were to validate the system's end-to-end performance in a real-world setting and to identify key considerations for future development.

Methodology

Data collection was repeated twice following a structured procedure. Initially, the ground-truth instruments, including the lux meter and luminance meter, were placed directly adjacent to the White Reference Target. The drone then executed a manual flight mission, which involved ascending vertically and sequentially to five distinct altitudes: 2 m, 3 m, 5 m, 10 m, and 15 m above the target. At each altitude, the drone hovered to capture a single, nadir-view JPEG aerial image while a corresponding reading was simultaneously recorded by the ground-truth instruments.

The experiment was conducted using the equipment detailed in Table 1, following the physical geometry illustrated in Figure 1.

Table 1 Employed Experimental Equipment

Role in Experiment	Equipment	Key Purpose
Aerial Platform	DJI M300 Drone	Provided a stable platform for aerial image capture.
Aerial Imaging	DJI Zenmuse H30T	Captured RGB images for illuminance calculation.
Ground-Truth Illuminance	Gigahertz Optik BTS256-EF Lux Meter	Served as the reference for ground-truth illuminance (lx) data.
Ground-Truth Luminance	TechnoTeam LMK4	Served as the reference for ground-truth luminance (cd/m^2) data.

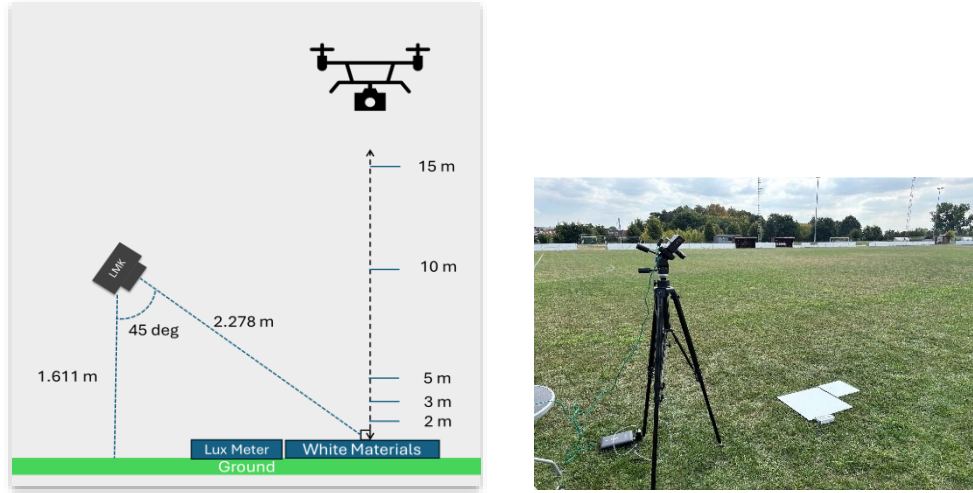


Figure 1 Equipment Setting and Geometry

Result

The results of the field validation experiment are summarised in Table 2 and Figures 2-4. The primary finding is that altitude and model factor have few effect to produce accurate absolute illuminance measurements when deployed in natural daylight.

A quantitative summary of the system's performance is presented in Table 2. The drone-based measurements exhibited a significant and systematic underestimation across the all tested conditions. The MAPE was approximately 99% for all factor groups, accompanied by a large negative mean bias of roughly -57,000 lx (predicted – ground-truth). The negative R^2 values further confirm a complete lack of correlation between the predicted and ground-truth illuminance values.

Table 2 Overall System Performance

Factor	MAPE (%)	Mean Bias (lx)	R^2
Assumed Area Reflectance	99.02266	-57276.9	-4.50849
Reference Target	99.89035	-57724.4	-4.58578
Reference Target 2	99.89877	-57728.8	-4.58653
	---	---	---
Altitude = 2m	99.62104	-66907.5	-4.15033
Altitude = 3m	99.62494	-67790.8	-4.31017
Altitude = 5m	99.62379	-63170	-5.45842
Altitude = 10m	99.61602	-50562.5	-20.4302
Altitude = 15m	99.53384	-39452.6	-49.1194
	---	---	---
Model = halogen	99.66864	-57610.2	-4.56603
Model = ledcube_cho	99.50999	-57529.4	-4.55229

Model = ledcube_ch2	99.39073	-57467.7	-4.54166
Model = ledcube_ch3	100.009	-57781.4	-4.59487
Model = mix	99.44127	-57494.8	-4.54648

The influence of the different experimental factors on this primary failure was negligible, visualised in the MAPE heatmap (Figure 2). There was no meaningful performance difference between the standard white tile (Reference Target 1), the larger matte tile (Reference Target 2), or the assumed reflectance for the grass surface. Similarly, neither flight altitude nor the specific training model used led to any practical improvement in accuracy.

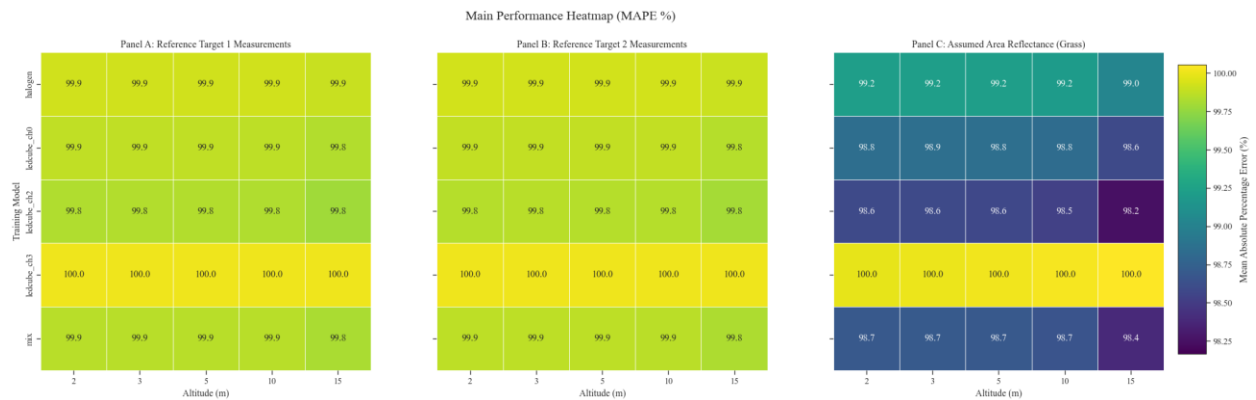


Figure 2 MAPE Analysis for Experimental Factors

The cause of the poor quantitative performance is diagnosed by the scatter plot in Figure 3. While the ground-truth illuminance values spanned a wide range (approx. 40,000 to 100,000 lx), the system's predicted values were consistently confirmed to a low range (0 to 1,000 lx). The data points form flat lines near the bottom of each plot, demonstrating a massive discrepancy between the predicted and actual illuminance.

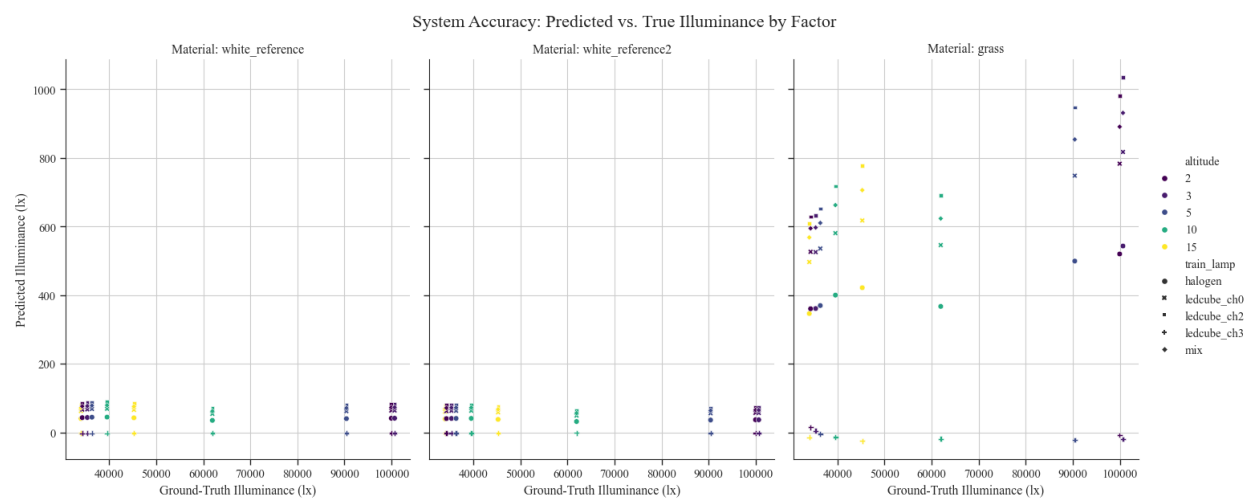


Figure 3 Prediction vs Ground-Truth Illuminance

To investigate the root cause of the system's poor performance, the camera's RGB channel response and calculated reflectance were analysed (Figure 4). The data points corresponding to both the standard white tile (white_reference) and the larger matte tile (white_reference 2) consistently cluster at the maximum normalised channel value of 1.0. This indicates that the camera sensor was completely saturated. In contrast, the grass surface shows a more linear response with its channel values remaining well below this saturation limit. In addition, the reflectance of the white materials, when calculated from the ground-truth luminance and illuminance meter readings, yielded physically impossible values far exceeding 1.0.

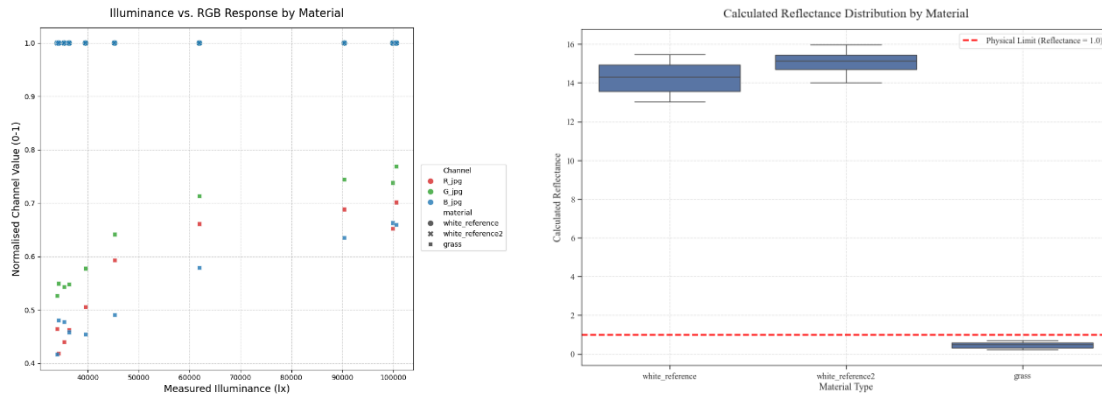


Figure 4 RGB Saturation and Invalid Reflectance

Discussion

The one finding of this study is the fundamental failure of the lab-trained models to generalise to the outdoor environment, resulting in a near-total error ($\text{MAPE} \approx 99\%$) and a severe underestimation of illuminance as we expected. From the previous experiment, the root causes could be the spectral mismatch between the daylight and the controlled laboratory illuminants.

As diagnosed in the results (Figure 4), the camera sensor and ground-truth luminance meter were heavily saturated by the high-reflectance white targets. This indicates that the light intensity in the field was orders of magnitude higher than what the system was calibrated for. The models, having been trained on a relatively low-intensity dataset, were operating far outside their valid input range. Consequently, their invalid RGB caused effectively luminance prediction failure leading to performance degradation (Figure 2, 3).

Within the context of this primary model failure, the secondary experimental factors—flight altitude and the choice of training model—had no meaningful effect on performance. The error introduced by the model's inability to handle daylight was so significant that it completely overwhelmed any subtle variations that might have been caused by changing the flight height or the specific lab illuminant used for training. This finding is critical, as it suggests that optimising operational parameters like altitude is a secondary concern. The fundamental calibration approach must be addressed before any meaningful analysis of flight parameters can be conducted.

While model saturation is the most evident cause of failure, it is also necessary to consider potential methodological limitations. One such factor is the synchronisation between the ground-truth instrument readings. The calculation of reflectance relies on the assumption that the luminance and the ground illuminance are measured at the exact same instant.

Any slight desynchronisation, even by a fraction of a second, could be significant under variable daylight conditions (e.g., subtle, fast-moving clouds). Such a timing mismatch could lead to the calculation of erroneous and physically impossible reflectance values, like those observed to be greater than 1.0 (Figure 4). While instrument saturation is the more dominant and likely cause of these invalid values, a lack of perfect synchronisation cannot be entirely ruled out as a contributing factor and should be a key consideration in the design of future experiments.

Future Work

The primary limitation identified in this study was the static nature of the lab-based calibration, which fails to account for the dynamic spectral characteristics of daylight. To overcome this, a promising direction for future work is the development of an adaptive, in-situ model selection system. This approach would move away from a "one-size-fits-all" model and instead dynamically choose the most appropriate model based on the real-time lighting conditions.

The proposed workflow would consist of the following key components:

1. **Develop a Comprehensive Model Library:** Instead of a single model, a library of specialised luminance prediction models would be pre-trained. Each model would be optimised for a specific, distinct illuminant (e.g., CIE D65 for direct sun, D55 for overcast sky, various artificial sources like high-pressure sodium, etc.). This creates a set of "expert" models for different lighting scenarios.
2. **Use a White Reference for In-situ Illuminant Characterisation:** A spectrally neutral white reference patch, placed in the field during the survey, would act as a real-time "spectral sensor." The linear RGB values extracted from this patch's ROI in an aerial image provide a unique chromatic signature of the prevailing ambient illuminant.
3. **Implement Automated Model Selection via Clustering:** An unsupervised clustering algorithm (e.g., k-means) would be used to automate the selection process.
 - **Setup:** The RGB signatures from the white patch under each illuminant in the model library would be clustered. Each resulting cluster centroid would represent a distinct "illuminant type" and be mapped to its corresponding expert model.
 - **Field Deployment:** During a flight, the live RGB signature from the white patch would be captured. The system would then find the nearest cluster centroid and dynamically select the associated model from the library to process the imagery from that entire survey.

This adaptive framework would represent a significant advancement. It directly addresses the spectral mismatch problem identified in this work by enabling the system to self-characterise the lighting environment and apply the most suitable calibration on the fly. This would dramatically

increase the system's robustness and potential for accurate illuminance measurements across a wide variety of real-world daylight conditions.

Conclusion

As anticipated, the lab-calibrated models performed poorly under daylight conditions. The central finding of this study is that operational factors like flight altitude and model choice did not mitigate this failure. The dominant error introduced by instrument saturation and spectral mismatch was so significant that it masked any potential influence from these secondary parameters. This result demonstrates that optimising flight procedures is ineffective until the fundamental, in-situ calibration problem is solved. Therefore, future work should prioritise adaptive calibration strategies such as those outlined above.