HDR Radiance Learning and Shadow Regularization for Satellite NeRF 3D Reconstruction

Yongjun Song, Pablo d’Angelo

German Aerospace Center (DLR), Earth Observation Center (EOC),
Münchener Str. 20, 82234 Oberpfaffenhofen, Germany
(yongjun.song@dlr.de, pablo.angelo@dlr.de)

Keywords: HDR Radiance Learning, Shadow Regularization, Neural Radiance Fields (NeRF), Satellite 3D Reconstruction.

Abstract

High dynamic range (HDR) variations in satellite optical imagery arise from extreme differences in surface reflectance and illumination conditions. Conventional satellite NeRF frameworks are typically trained on tone-mapped or radiometrically enhanced images, where nonlinear preprocessing alters the physical relationship between measured pixel values and true scene radiance. This leads to biased photometric optimization and loss of geometric fidelity, especially under strong illumination contrasts. To address these limitations, we propose an HDR-consistent learning framework that integrates RawNeRF-style radiance supervision with shadow regularization. The method trains directly on raw satellite imagery using a logarithmic, tone mapping–aware loss that preserves linear radiance and stabilizes optimization under high dynamic range conditions. In parallel, a soft shadow regularization constrains network-predicted shadows using geometric cues derived from solar ray casting, promoting physically consistent irradiance decomposition. Experiments on four AOIs from the DFC2019 dataset demonstrate that HDR-aware radiance learning significantly improves DSM accuracy by maintaining linear radiometric consistency. The proposed shadow regularization also improves geometric consistency in structure-dominated urban scenes, although its effect is limited in vegetation-dominant areas where shadow cues are less informative. Although performance gains are smaller in vegetation-dominant areas, the results confirm that combining HDR radiance learning with geometric shadow regularization yields more radiometrically consistent and geometrically accurate 3D reconstruction from satellite imagery.

1. Introduction

Satellite optical imagery inherently exhibits a high dynamic range (HDR) due to extreme variations in surface reflectance and atmospheric conditions. To improve visual interpretability and inter-scene consistency, radiometric enhancement procedures such as outlier removal and tone mapping are commonly applied. These processes suppress sensor-specific noise (e.g., dark pixels and saturated highlights) and expand information in low-intensity regions. However, such enhancements alter the original radiometric distribution of the captured signal, especially when values are clipped or compressed by tone mapping curves and subsequent quantization.

Recent advances in rendering-based 3D reconstruction, including Neural Radiance Fields (NeRF) (Mildenhall et al., 2021), learn to represent scene geometry and appearance by minimizing the discrepancy between rendered radiance and observed RGB values through differentiable volume rendering. Accurate learning therefore requires input data that preserve the physical radiance measured by the sensor. However, when radiometric enhancement is applied, details in both bright and dark regions are lost, and the nonlinear tone mapping function distorts the correspondence between predicted and true radiance. To address this issue, RawNeRF (Mildenhall et al., 2022) proposed training directly on raw HDR imagery using a tone mapping–aware weighted loss, enabling stable optimization while maintaining the full dynamic range of scene radiance.

In the satellite domain, NeRF-based methods such as Sat-NeRF (Marí et al., 2022) and EONeRF (Marí et al., 2023) extend this framework by explicitly modeling irradiance components—ambient illumination, albedo, and shadows—to account for complex lighting conditions in multi-date imagery. Shadows are typically represented either through network-predicted components or by casting secondary rays along the solar direction. Existing approaches have focused on radiometrically enhanced datasets (e.g., tone-mapped imagery) where the radiance distribution has already been compressed for visual consistency. Within this setting, several works have introduced physically motivated formulations to better explain shadow phenomena. S-NeRF (Derksen and Izzo, 2021) proposed a solar-correction loss that regularizes predicted shadows using transmittance. EONeRF further refined this idea by computing geometric shadows directly from reconstructed 3D geometry, improving altitude accuracy and generalization to novel solar directions. SUNDIAL (Behari et al., 2024) extended the framework by jointly estimating the sun direction and decomposing direct, ambient, and complex illumination. More recently, S-EO (Masquil et al., 2025) leveraged externally derived shadow masks as additional supervision cues, demonstrating improved 3D reconstruction under few-shot conditions.

However, none of these methods directly address the radiometric instability that arises when learning from raw HDR satellite imagery. In the absence of tone mapping, shaded regions exhibit extremely narrow dynamic ranges, making shadow estimation noisy and ambiguous. Geometric shadows derived purely from transmittance also fail to reproduce realistic intra-shadow variations caused by interreflection or atmospheric scattering. These limitations motivate a self-supervised approach that combines the physical interpretability of geometric shadows with the flexibility of learnable, network-predicted shadows.

These limitations highlight the need for a unified framework that can preserve HDR radiance while maintaining physically consistent shadow modeling. Motivated by these challenges, this study explores two complementary directions: (1) HDR-

This contribution has been peer-reviewed. The double-blind peer-review was conducted on the basis of the full paper.
https://doi.org/10.5194/isprs-annals-XI-2-2026-341-2026

consistent radiance learning for robust NeRF-based reconstruction from raw satellite imagery, and (2) shadow regularization, where physically derived geometric shadows serve as weak supervision for network-based shadow estimation. Together, these strategies aim to improve both radiometric fidelity and geometric accuracy in 3D reconstruction from satellite imagery.

2. Methods

2.1 HDR Radiance Learning with RawNeRF Loss

NeRF (Mildenhall et al., 2021) represents a scene as a continuous volumetric function that maps 3D coordinates and view directions to emitted radiance and density. The radiance field is optimized by minimizing the photometric error between rendered and observed pixel colors through differentiable volume rendering. This formulation assumes that the training images provide linear radiance values, allowing the network to learn the physical light transport across viewpoints. However, most real-world and satellite imagery undergo several preprocessing steps such as nonlinear tone mapping, quantization, and dynamic range compression. These preprocessing steps distort the relationship between measured pixel values and true scene radiance.

RawNeRF (Mildenhall et al., 2022) was proposed to address these issues by enabling NeRF to learn directly from linear HDR raw data, rather than preprocessed tone-mapped images. This design allows the network to recover true scene radiance consistent with physical light transport, while explicitly accounting for the nonlinear camera response. However, when applying a conventional photometric loss to tone-mapped images, the nonlinear transformation introduces bias that distorts the supervision signal. A naïve tone-mapped mean-squared error (MSE) loss,

$$\mathcal{L}_{\psi}(\hat{y}, y) = \sum_{i} \|\psi(\hat{y}_i) - \psi(y_i)\|^2, \tag{1}$$

produces biased gradients when the raw observation y is corrupted by zero-mean noise, since a nonlinear tone mapping function $\psi(\cdot)$ changes the expected value ($\mathbb{E}[\psi(y)] \neq \psi(\mathbb{E}[y])$). To avoid this bias, RawNeRF replaces the explicit tone mapping with a weighted MSE loss:

$$\tilde{\mathcal{L}}_{\psi}(\hat{y}, y) = \sum_{i} [\psi'(\mathrm{sg}(\hat{y}_i))]^2 \, (\hat{y}_i - y_i)^2, \tag{2}$$

where $\psi'(\cdot)$ is the derivative of the tone mapping curve and $\mathrm{sg}(\cdot)$ denotes a stop-gradient operation. This formulation rebalances the loss magnitude across intensity levels, suppressing dominance by bright regions and improving contrast reconstruction in darker areas.

We follow RawNeRF and model the tone mapping function as a logarithmic curve,

$$\psi(x) = \log(x + \epsilon), \quad \psi'(x) = \frac{1}{x + \epsilon}, \tag{3}$$

where $\epsilon = 10^{-6}$ ensures numerical stability near zero intensity. The resulting log-based weighted MSE loss preserves the statistical unbiasedness of HDR supervision. In addition, we evaluate an alternative tone mapping model commonly used in satellite imaging, defined by the gamma correction curve (RawNeRF-γ):

$$\psi(x) = x^{\frac{1}{\gamma}}, \quad \psi'(x) = \frac{1}{\gamma} x^{\frac{1}{\gamma} - 1}, \tag{4}$$

where $\gamma = 2.2$ amplifies gradients in dark regions and attenuates them in bright regions. This HDR-aware formulation provides a physically consistent basis for radiance learning from raw satellite imagery.

2.2 Shadow–Irradiance Color Modeling

Following EONeRF (Marí et al., 2023), we model the observed pixel color as the combined effect of surface albedo, illumination, and per-image radiometric correction. For a camera ray r, the rendered color is expressed as:

$$c(r) = A_j \Big(\ell(r) \sum_{i=1}^{N} T_i \alpha_i c_{a,i} \Big) + b_j, \tag{5}$$

where A_j and b_j are per-image affine parameters compensating for cross-date radiometric bias, and $c_{a,i}$ is the intrinsic albedo at each sampled 3D point. The opacity α_i and transmittance T_i are computed from the predicted volume density σ_i as:

$$\alpha_i = 1 - \exp(-\sigma_i \, \delta_i), \qquad T_i = \prod_{j<i} (1 - \alpha_j), \tag{6}$$

where δ_i denotes the sampling interval along the ray.

The irradiance term $\ell(r)$ represents the mixture of direct and ambient illumination:

$$\ell(r) = s(r) + \big(1 - s(r)\big)\mathbf{a}, \tag{7}$$

where $\mathbf{a}$ is the global ambient color and $s(r) \in [0, 1]$ indicates the proportion of direct illumination—low values correspond to shaded regions, and high values to fully lit areas.

While EONeRF computes geometric shadows $s_{\mathrm{geo}}(x)$ via solar ray casting, such transmittance-based values cannot fully describe real shadow phenomena in satellite imagery, where reflected and scattered light produce spatially varying shadow intensities. To better capture these effects, we replace $s_{\mathrm{geo}}(x)$ with a learnable network-predicted shadow $s_{\mathrm{net}}(r)$, similar to SatNeRF (Marí et al., 2022), allowing data-driven modeling of illumination and shadow interactions. The overall illumination factor becomes

$$s(r) = \tau(r) \, s_{\mathrm{net}}(r), \tag{8}$$

where $\tau(r)$ is an arbitrary transient scalar that models inter-date transient effects.

In this study, we use the network architecture illustrated in Figure 1, which extends EONeRF by adding a shadow-prediction branch that outputs $s_{\mathrm{net}}(r)$ from volumetric features and the solar direction.

2.3 Shadow Regularization

EONeRF explicitly models solar illumination by casting secondary rays toward the sun, producing a geometric shadow that depends on the transmittance along each solar ray. While this physically grounded formulation ensures a direct link between geometry and shading, the resulting shadow values are inherently binary—determined solely by whether a surface occludes the solar path. Consequently, it cannot represent the continuous shadow intensities observed in real satellite imagery, where

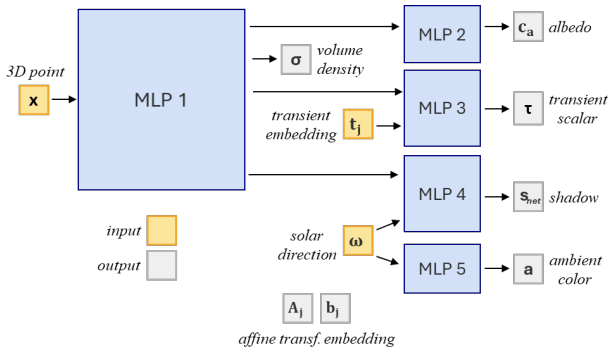

Figure 1. Modified EO-NeRF network architecture.

reflected light and atmospheric scattering produce gradual variations within shadowed regions.

S-EO (Masquil et al., 2025) proposed a complementary shadow regularization strategy that leverages externally derived shadow masks to improve geometry optimization during EONeRF training. By enforcing consistency between rendered shadow and shadow masks, the method enhanced 3D reconstruction quality, as demonstrated by improved DSM accuracy across varying numbers of input views. Importantly, this approach supplements the geometric model rather than replacing it, providing additional supervision to guide shadow learning.

In our formulation, shadows are predicted by a learnable network rather than derived geometrically. Without constraints, this can lead to ambiguity between albedo and shading. To mitigate this, we use the transmittance-based geometric shadow $s_{\text{geo}}(r)$ obtained by solar ray casting as a weak prior. Specifically, $s_{\text{geo}}(r)$ is thresholded to form a pseudo binary mask:

$$M_{\text{geo}}(r) = \mathbf{1}(s_{\text{geo}}(r) < 0.5)\,, \tag{9}$$

which regularizes the network-predicted shadow $s_{\text{net}}(r)$ through a soft regression loss:

$$\mathcal{L}_{\text{reg}} = \mathbb{E}_r\left[M_{\text{geo}} s_{\text{net}}^2 + (1 - M_{\text{geo}})(1 - s_{\text{net}})^2\right]. \tag{10}$$

This regularization encourages low s_{net} values in geometrically shaded regions and high values elsewhere, guiding the network toward physically consistent and spatially adaptive shadow estimation. Unlike purely geometric transmittance, this hybrid approach maintains geometric plausibility while allowing realistic shadow variability, resulting in more faithful radiance reconstruction in complex urban scenes. To reduce error propagation from inaccurate early-stage geometry, all shadow-related supervision is activated only from epoch 2. After warm-up, the shadow-prior term is computed only in potentially shadowed regions ($s_{\text{net}} \leq 0.5$).

It should be noted that the proposed shadow regularization is not intended as an independent supervisory signal. The geometric shadow $s_{\text{geo}}(r)$, derived from transmittance, provides a simplified and inherently binary approximation that is sensitive to geometric inaccuracies and does not capture indirect illumination effects. As a result, enforcing it in isolation may bias the optimization. In this work, shadow regularization is used as a weak constraint under HDR-consistent radiance learning. In combination with RawNeRF supervision, it helps reduce ambiguity in the decomposition of albedo and shading while allowing continuous shadow variations.

3. Data and Experimental Setup

3.1 Datasets

We evaluate our method using the public *2019 IEEE GRSS Data Fusion Contest* (DFC2019) dataset (Bosch et al., 2019), which provides multi-date WorldView-3 (WV-3) imagery over Jacksonville, Florida. Four Areas of Interest (AOIs)—004, 068, 214, and 260—are selected for evaluation. Each AOI covers approximately 256×256 m^2 and consists of 10–20 multi-date crops of WV-3 images with a ground sampling distance of 30–50 cm at nadir. The DFC2019 Track 3 provides both WV-3 multispectral (MSI) products with only pansharpening applied (16-bit depth) and radiometrically processed 8-bit RGB products (including operations such as outlier removal and gamma correction), together with lidar-derived digital surface models (DSMs) and corresponding Rational Polynomial Camera (RPC) models. Most prior satellite NeRF studies on DFC2019 use the provided RGB products for training. In contrast, to support HDR radiance learning, we use MSI products. Since WV-3 collects raw sensor data with an 11-bit dynamic range, each image is normalized by dividing pixel values by $(2^{11} - 1)$ to preserve the linear radiometric scale. This configuration effectively corresponds to learning from raw HDR imagery.

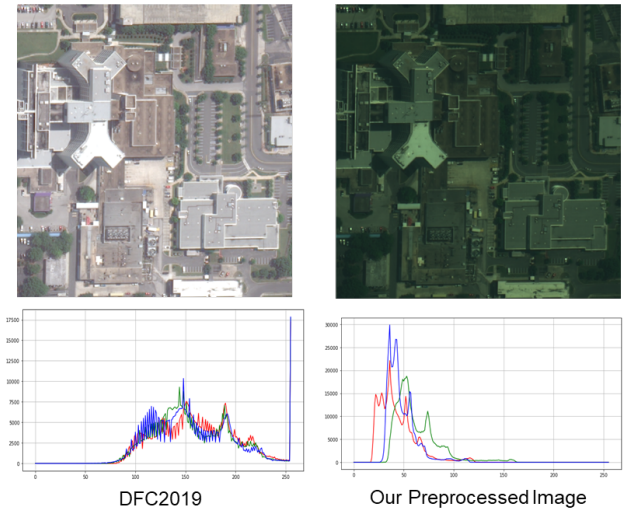

Figure 2. Comparison of input radiometric distributions: (left) DFC2019 image, (right) preprocessed image used in our experiment with RGB histograms. For direct comparison with the 8-bit DFC2019 scale, the histogram x-axis of the preprocessed image is multiplied by 255.

Figure 2 compares the commonly used DFC2019 RGB image and intensity histogram with our preprocessed image and histogram. The preprocessed input shows a clear channel-wise redistribution of intensities, indicating that radiometric transformations alter the original linear response and motivating HDR-consistent training on linear radiance. For geometric consistency, bundle adjustment (Marí et al., 2021) is applied to refine the RPC parameters, and each image is cropped to spatially correspond with the lidar-derived DSMs for evaluation. Quantitative comparisons are performed against the lidar DSMs using the mean absolute error (MAE) in meters.

This contribution has been peer-reviewed. The double-blind peer-review was conducted on the basis of the full paper.
https://doi.org/10.5194/isprs-annals-XI-2-2026-341-2026 | © Author(s) 2026. CC BY 4.0 License.

3.2 Experimental Setup

We validate the proposed components through controlled ablation studies on two satellite NeRF frameworks, SatNeRF (Marí et al., 2022) and EONeRF (Marí et al., 2023). These baselines differ in irradiance modeling and network structure, enabling evaluation of the generality of HDR-aware radiance learning and shadow regularization. All experiments are conducted on four AOIs from the DFC2019 dataset using identical camera poses, cropped images, and optimization settings for fair comparison.

(a) RawNeRF Loss. We evaluate four loss variants to analyze the effect of HDR radiance learning: (i) standard photometric MSE, (ii) MSE-γ defined in Eq. (1), where both rendered and reference RGBs are tone-mapped with a standard γ function before error computation, (iii) RawNeRF-γ using the gamma-based weighting of Eq. (4), and (iv) RawNeRF employing the log-based weighting of Eq. (3). This comparison isolates the impact of tone-mapping–aware weighting on learning radiometrically linear radiance from satellite imagery.

(b) Shadow Regularization. The effect of shadow regularization is examined using the modified EONeRF model (EONeRF′, Figure 1). We test four configurations: without RawNeRF loss or regularization, with RawNeRF loss only, with regularization only, and with both components combined. This ablation clarifies the complementary roles of HDR radiance supervision and geometric priors in improving 3D reconstruction fidelity. The overall loss function is given by

$$\mathcal{L}_{\text{total}} = \tilde{\mathcal{L}}_{\psi} + \lambda_{\text{reg}}\,\mathcal{L}_{\text{reg}}, \tag{11}$$

where the regularization weight λ_{reg} is set to 0.01 when used with the RawNeRF loss and 0.0005 otherwise. This accounts for the change in loss scale introduced by the RawNeRF formulation, which increases the magnitude of the photometric term for normalized inputs in [0,1].

4. Experimental Results and Discussion

4.1 Quantitative Evaluation

Table 1 shows that applying the HDR-aware RawNeRF loss consistently improves DSM accuracy for both SatNeRF and EONeRF. SatNeRF benefits most from the γ-based weighting, while EONeRF performs best with the logarithmic variant, especially on AOIs (068, 214, and 260) that contain dense urban structures and strong shadow variations. These results confirm that HDR-aware weighting improves model robustness to large radiometric variations, resulting in more accurate DSM reconstruction.

Table 2 further examines the effect of shadow regularization. When used alone, the regularization increases MAE due to unstable shadow predictions. This behavior can be attributed to the nature of the geometric shadow prior, which is derived from transmittance and thus provides a simplified and inherently binary approximation of shadow formation. As a result, it is sensitive to inaccuracies in early-stage geometry and may introduce bias when applied in isolation. However, when combined with the RawNeRF loss, EONeRF′ achieves the lowest errors, particularly in complex urban areas (068 and 214). This suggests that geometric-shadow guidance is effective only when HDR-consistent radiance learning is present, where the radiometric supervision stabilizes the optimization. Under this condition, shadow regularization provides additional geometric guidance and improves overall DSM quality. In contrast, areas with extensive vegetation (004) show no benefit from shadow regularization, implying that the method is less effective in vegetation-dominant scenes where shadow cues are weak or diffuse. Notably, unlike the original EONeRF, which employs an uncertainty-based loss (Martin-Brualla et al., 2021) to handle transient objects, our model does not include such a term. Nevertheless, the performance gap remains small, and in some AOIs our method even achieves lower errors, confirming the robustness of the proposed formulation without explicit transient modeling.

Model	Loss Type	MAE [m] ↓			
		004	068	214	260
SatNeRF	MSE	2.26	2.28	4.46	2.88
	MSE-γ	2.31	2.01	4.11	2.68
	RawNeRF-γ	2.31	**2.00**	**3.97**	**2.51**
	RawNeRF	**2.23**	2.10	3.98	2.87
EONeRF	MSE	1.51	1.36	3.05	2.58
	MSE-γ	**1.44**	1.27	2.44	1.78
	RawNeRF-γ	1.46	1.22	2.51	1.99
	RawNeRF	1.64	**1.12**	**1.93**	**1.48**

Table 1. Quantitative results of HDR radiance learning under different loss variants.

Model / Variant	MAE [m] ↓			
	004	068	214	260
SatNeRF†	1.37	1.14	2.54	2.31
EONeRF†	**1.29**	1.11	1.81	**1.38**
EONeRF′ (w.o.)	1.47	1.16	2.23	1.93
EONeRF′ (RawNeRF)	1.38	1.07	1.79	1.44
EONeRF′ (reg)	2.33	1.17	2.22	1.98
EONeRF′ (RawNeRF+reg)	1.55	**1.05**	**1.78**	1.46

Table 2. Quantitative results of shadow regularization and HDR radiance learning on four AOIs.[1]

To further analyze the proposed components, we evaluate reconstruction accuracy across object classes (ground, trees, and buildings). Table 3 reports object-wise MAE for two urban scenes (AOIs 068 and 214), where shadow cues are most informative. The results show that the impact of shadow regularization varies across surface types. Ground and vegetation (trees) regions exhibit improvements, suggesting that shadow cues provide complementary geometric information in areas with weaker radiometric signals. In contrast, the effect on building regions is marginal, as their geometry is already well constrained by multi-view consistency and strong radiometric cues. Overall, the contribution of shadow regularization is spatially non-uniform and becomes more relevant where radiometric information is limited. Consequently, global metrics such as MAE may not fully reflect these localized effects.

4.2 Qualitative Analysis

To better understand the role of shadow modeling, Figure 3 contrasts the geometric shadow from solar ray casting with the network-predicted shadow produced by our model. The geometric shadow represents binary occlusion—regions

[1] The SatNeRF† and EONeRF† results were reproduced using the official datasets and configurations described in the original SatNeRF and EONeRF papers

Model / Variant	MAE [m] ↓			
	Ground	Trees	Buildings	Overall
EONeRF′ (w.o.)	1.57	3.47	1.37	1.69
EONeRF′ (RawNeRF)	1.27	2.80	**1.20**	1.42
EONeRF′ (reg)	1.60	3.27	1.37	1.68
EONeRF′ (RawNeRF+reg)	**1.24**	**2.71**	1.22	**1.41**

Table 3. Object-wise DSM accuracy for ground, vegetation (trees), and buildings on two AOIs (068 and 214).

are either illuminated or fully dark—whereas the network-predicted shadow displays smooth gradations that reflect indirect illumination and scattering effects. This behaviour emphasizes the necessity of a learnable shadow representation to model realistic lighting conditions in satellite imagery.

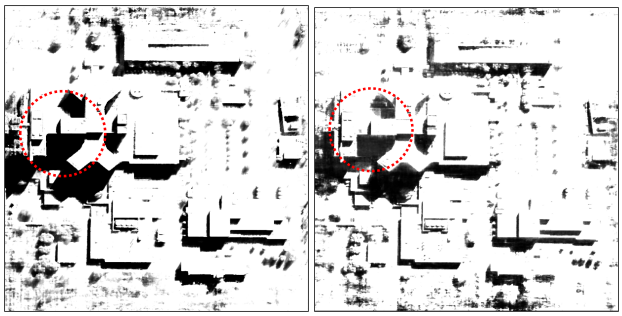

Figure 3. Comparison between (a) geometric shadow and (b) network-predicted shadow with regularization, both visualized with intensity stretching for enhanced contrast.

Next, the effect of HDR-aware loss formulations on reconstruction quality is illustrated in Figure 4. For the urban AOIs (068, 214, and 260), applying the RawNeRF loss noticeably sharpens building edges and improves structural separability. The log-based weighting produces the clearest object boundaries and more stable elevation gradients, while the γ-based loss achieves similar yet slightly less pronounced results. These findings confirm that HDR-consistent radiance learning preserves subtle radiometric cues that support more accurate geometric reconstruction. In contrast, performance in the vegetation-dominant scene (004) shows little improvement, indicating that tone-mapping–aware weighting is less effective when shading information is weak or spatially diffuse.

The impact of the proposed shadow regularization is examined in Figure 5. When comparing EONeRF† and EONeRF′ (RawNeRF+reg), both achieve similar DSM quality across the urban scenes, reconstructing consistent building geometries and elevation patterns. Despite excluding transient-object modeling, EONeRF′ maintains comparable accuracy, suggesting that HDR-consistent learning and geometric shadow regularization together ensure stable reconstruction across varying illumination conditions. However, the 004 scene again exhibits irregular artifacts—particularly along flat road surfaces—indicating that in vegetation-rich regions with limited shading cues, geometric priors offer less effective guidance.

Finally, Figure 6 presents the difference maps between reconstructed DSMs and reference lidar data for EONeRF† and EONeRF′. In the urban AOIs (068, 214, 260), both models display similar residual patterns, with errors mainly concentrated along building boundaries and rooflines. Although EONeRF′ tends to yield smoother residuals near shadowed façades, its errors are slightly more spatially dispersed than

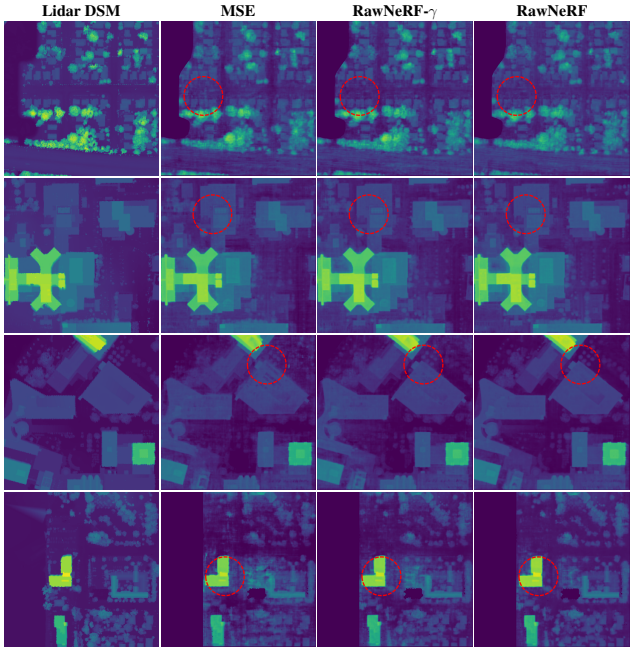

Figure 4. DSM reconstruction results using EO-NeRF across four AOIs (004, 068, 214, and 260). Columns show (from left to right): Lidar DSM, without RawNeRF, RawNeRF-γ, and RawNeRF. Rows correspond to AOIs from top to bottom.

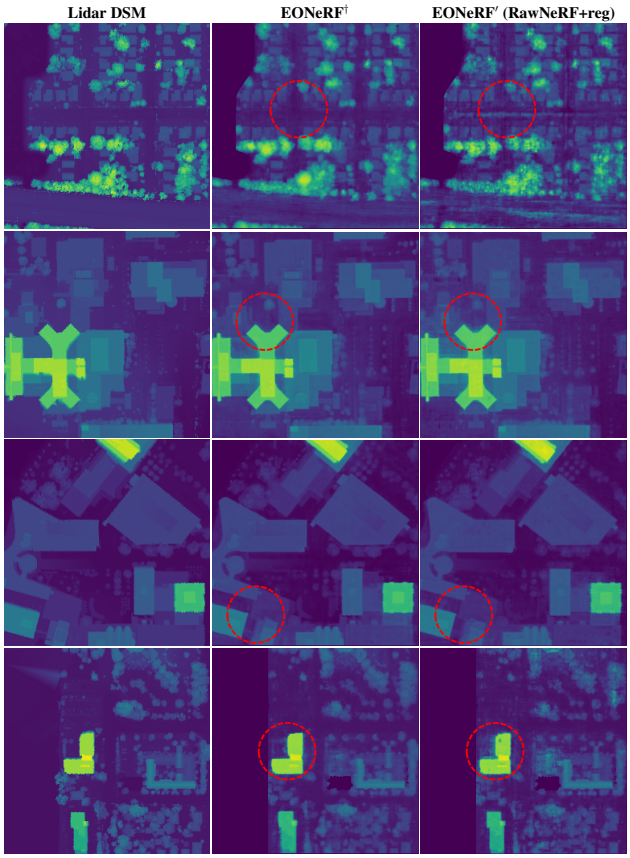

Figure 5. DSM reconstruction results across four AOIs (004, 068, 214, and 260). Columns show (from left to right): Lidar DSM, EO-NeRF†, and EO-NeRF′ with RawNeRF loss and shadow regularization. Rows correspond to AOIs from top to bottom.

This contribution has been peer-reviewed. The double-blind peer-review was conducted on the basis of the full paper.
https://doi.org/10.5194/isprs-annals-XI-2-2026-341-2026 | © Author(s) 2026. CC BY 4.0 License.

those of EONeRF†, implying that quantitative HDR gains do not always correspond to locally uniform geometry. For the vegetation-dominant AOI (004), large fluctuations appear over roads and vegetation for both methods, underscoring the inherent difficulty of reconstructing low-contrast surfaces. Overall, while HDR-aware learning improves general reconstruction accuracy, localized geometric errors remain in areas lacking distinct shading or structure.

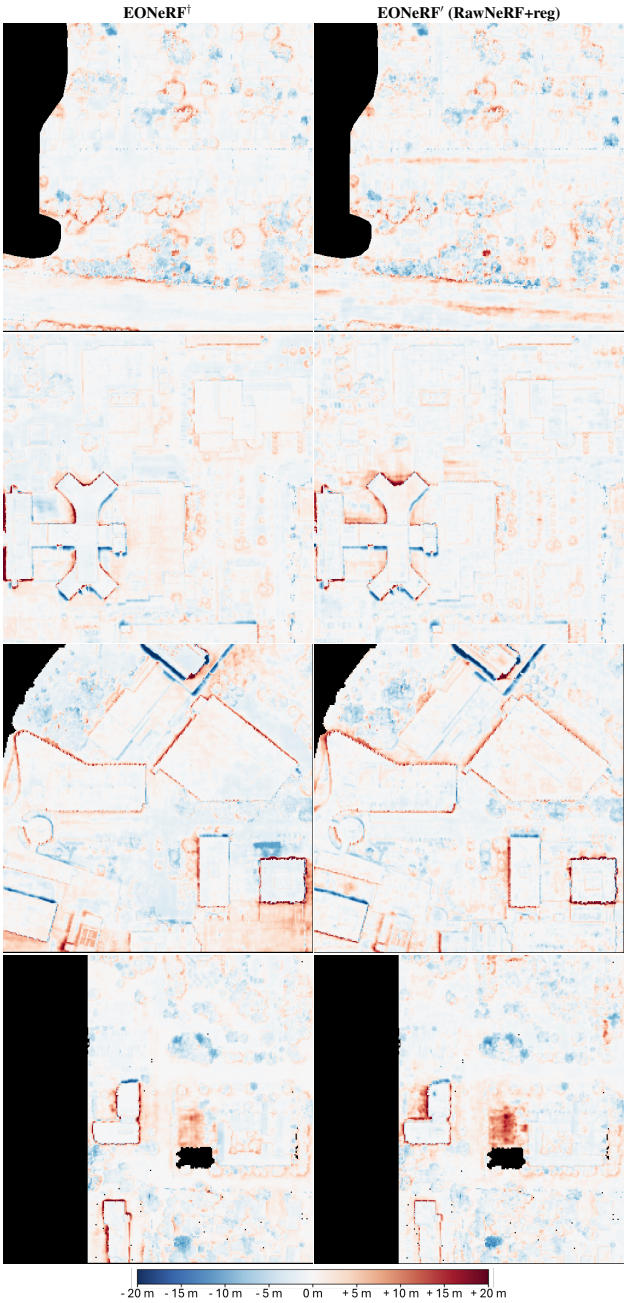

Figure 6. Difference maps between reconstructed DSMs and reference Lidar DSM across four AOIs (004, 068, 214, and 260). Columns show (from left to right): EO-NeRF† and EO-NeRF′ (RawNeRF+reg). A common color scale is shown below.

5. Conclusion

This study investigated HDR radiance learning and shadow regularization for satellite NeRF-based 3D reconstruction using minimally preprocessed WV-3 imagery. By applying the RawNeRF loss, the network was trained directly on radiometrically linear data, preserving the true HDR response of the satellite sensor and maintaining consistency with the physically inspired image formation model. The proposed shadow regularization further guided the learning of network-predicted shadows through geometric cues, improving shadow estimation and enhancing geometric reconstruction quality in dense urban areas. By avoiding heuristic preprocessing and tone-mapping procedures, the modeling pipeline better aligns with the physical image rendering equation and provides additional robustness in cases where radiometric corrections—often based on quantile stretching—fail, such as under partial cloud cover or strong illumination contrast.

Nevertheless, the method shows limited effectiveness in vegetation-dominated or suburban scenes, where shadow cues contribute less to geometric reconstruction. In such regions, shadows tend to be weaker, more spatially diffuse, and less directly correlated with stable geometry due to volumetric effects and temporal variability. As a result, both geometric shadow priors and learned shadow representations provide limited constraints for geometry reconstruction. This behavior is consistent with the results observed in AOI 004 and reflects a broader challenge in satellite-based neural reconstruction. Moreover, similar to previous network-based shadow models, the learned shadows occasionally introduce uneven surfaces compared to the smoother geometry reconstructed by geometric-shadow-based methods such as EONeRF.

Future work will aim to address these limitations by improving robustness across a broader range of landscapes and imaging conditions, including additional satellite sensors beyond WV-3. In particular, exploring alternative geometric priors or hybrid representations that better capture volumetric and low-contrast structures, such as vegetation, may further improve reconstruction quality. Extending the framework to other rendering methods, such as 3D Gaussian Splatting, could enable large-scale and cost-effective testing. Furthermore, developing a systematic approach for jointly modeling correlated rendering parameters—such as albedo, irradiance, and shadow components—may further improve physical interpretability and reconstruction stability in complex satellite observation environments.

References

Behari, N., Dave, A., Tiwary, K., Yang, W., Raskar, R., 2024. Sundial: 3d satellite understanding through direct ambient and complex lighting decomposition. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 522–532.

Bosch, M., Foster, K., Christie, G., Wang, S., Hager, G. D., Brown, M., 2019. Semantic stereo for incidental satellite images. *2019 IEEE Winter Conference on Applications of Computer Vision (WACV)*, IEEE, 1524–1532.

Derksen, D., Izzo, D., 2021. Shadow neural radiance fields for multi-view satellite photogrammetry. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 1152–1161.

Marí, R., de Franchis, C., Meinhardt-Llopis, E., Anger, J., Facciolo, G., 2021. A generic bundle adjustment methodology for indirect RPC model refinement of satellite imagery. *Image Processing On Line*, 11, 344–373.

Marí, R., Facciolo, G., Ehret, T., 2022. Sat-nerf: Learning multi-view satellite photogrammetry with transient objects and shadow modeling using rpc cameras. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 1311–1321.

Marí, R., Facciolo, G., Ehret, T., 2023. Multi-date earth observation nerf: The detail is in the shadows. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2035–2045.

Martin-Brualla, R., Radwan, N., Sajjadi, M. S., Barron, J. T., Dosovitskiy, A., Duckworth, D., 2021. Nerf in the wild: Neural radiance fields for unconstrained photo collections. *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 7210–7219.

Masquil, E., Marí, R., Ehret, T., Meinhardt-Llopis, E., Musé, P., Facciolo, G., 2025. S-eo: A large-scale dataset for geometry-aware shadow detection in remote sensing applications. *Proceedings of the Computer Vision and Pattern Recognition Conference*, 2383–2393.

Mildenhall, B., Hedman, P., Martin-Brualla, R., Srinivasan, P. P., Barron, J. T., 2022. Nerf in the dark: High dynamic range view synthesis from noisy raw images. *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 16190–16199.

Mildenhall, B., Srinivasan, P. P., Tancik, M., Barron, J. T., Ramamoorthi, R., Ng, R., 2021. Nerf: Representing scenes as neural radiance fields for view synthesis. *Communications of the ACM*, 65(1), 99–106.